A BOUNDARY VALUE PROBLEM FOR A SYSTEM OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS OF THE FIRST ORDER*

BY
GILBERT AMES BLISS

The boundary value problem to be considered in this paper is that of finding solutions of the system of differential equations and boundary conditions

$$(1)\qquad \begin{gathered} \frac{dy_i}{dx} = \sum_{\alpha=1}^{n} [A_{i\alpha}(x) + \lambda B_{i\alpha}(x)] y_\alpha(x), \\ \sum_{\alpha=1}^{n} [M_{i\alpha} y_\alpha(a) + N_{i\alpha} y_\alpha(b)] = 0 \qquad (i = 1, 2, \cdots, n). \end{gathered}$$

Such systems have been studied by a number of writers whose papers are cited in a list at the close of this memoir. The further details of references incompletely given in the footnotes or text of the following pages will be found there in full.

In 1909 Bounitzky defined for the first time a boundary value problem adjoint to the one described above, and discussed its relationships with the original problem. He constructed the Green's matrices for the two problems, and secured expansion theorems by considering the system of linear integral equations, each in one unknown function, whose kernels are the elements of the Green's matrix. In 1918 Hildebrandt, following the methods of E. H. Moore's general analysis, formulated a very general boundary value problem containing the one above as a special case, and established a number of fundamental theorems. In 1921 W. A. Hurwitz studied the more special system

$$(2)\qquad \begin{gathered} \frac{du}{dx} = [a(x) + \lambda] v(x), \qquad \frac{dv}{dx} = -\,[b(x) + \lambda] u(x), \\ \alpha_0 u(0) + \beta_0 v(0) = 0, \qquad \alpha_1 u(1) + \beta_1 v(1) = 0 \end{gathered}$$

and its expansion theorems, by the method of asymptotic expansions, and in 1922 Camp extended his results to a case where the boundary conditions have a less special form. Meanwhile Schur in 1921 published very general expansion theorems for the original system (1) under the hypothesis that the matrix of functions $B_{ik}(x)$ has all elements identically zero except those

* Presented to the Society, December 30, 1924; received by the editors in December, 1925.

in the principal diagonal, which are real, distinct, and positive for every value of x. The cases of Hurwitz and Camp are not included under this one even after a linear transformation. Carmichael exhibited in 1921–22 the analogies between certain algebraic problems and boundary value problems of many types, including those of the type (1), but without giving the details of the theory of the boundary value problems themselves. In 1923 Birkhoff and Langer showed that the large class of systems (1) for which the determinant of the functions $B_{ik}(x)$ is different from zero can be reduced, by a linear transformation of the functions $y_k(x)$ whose coefficients may be imaginary, to the simpler form for which all the functions $B_{ik}(x)$ with $i \neq k$ vanish identically. They obtained expansion theorems under the further hypotheses that the functions $B_{ii}(x)$ are all distinct from each other and from zero, though not necessarily real and positive as in Schur's paper, and that they satisfy certain other more artificial restrictions.* The problems considered by Hurwitz, Camp, and Schur are among those included in their theory. An abstract of the paper of Birkhoff and Langer was printed in 1922.†

The methods of Hurwitz, Camp, Schur, and Birkhoff and Langer are those of asymptotic expansions which for the more general cases become very complicated. In the following pages it will be shown that a large class of so called self-adjoint boundary value problems, analogous to equations with symmetric kernels in linear integral equation theory, can be treated by much more elementary methods.

A boundary value problem adjoint to the problem (1) has the form

$$\begin{aligned} &\frac{dz_i}{dx} = -\sum_{\alpha=1}^{n} (A_{\alpha i} + \lambda B_{\alpha i}) z_\alpha, \\ (3)\qquad &\sum_{\alpha=1}^{n} [P_{\alpha i} z_\alpha(a) + Q_{\alpha i} z_\alpha(b)] = 0 \qquad (i = 1, 2, \cdots, n), \end{aligned}$$

where the coefficients P_{ik}, Q_{ik} satisfy the relations

$$M_{i\alpha} P_{\alpha k} - N_{i\alpha} Q_{\alpha k} = 0 \qquad (i, k = 1, 2, \cdots, n).$$

The original problem is defined in this paper to be self-adjoint if the systems (1) and (3) are equivalent under a transformation of the form

$$z_i = \sum_{\alpha=1}^{n} T_{i\alpha}(x) y_\alpha,$$

* See pp. 83, 89, 109. One should note also the hypotheses on the roots of $D(\lambda)$ on pp. 98 and 105.

† See also an earlier abstract by Birkhoff, Bulletin of the American Mathematical Society, vol. 25 (1919), p. 442.

and it is definitely self-adjoint when a further hypothesis, explained in Section 2, is added. For a definitely self-adjoint boundary value problem the characteristic constants and functions are not only denumerably infinite in number, but the constants are real and each has its index equal to its multiplicity as a root of the characteristic determinant. The characteristic functions may be chosen real. Furthermore expansion theorems of a very general sort may be justified by relatively simple methods analogous to those of integral equation theory.

The problems of Hurwitz and Camp described above are definitely self-adjoint according to the definition of this paper. Those of Schur are never self-adjoint, and a sub-class only of the problems considered by Birkhoff and Langer have this property. The problems of Birkhoff and Langer are furthermore never definitely self-adjoint when the coefficients in the differential equations are real. On the other hand, the theory of definitely self-adjoint boundary value problems as presented here includes a large category of problems for which the determinant $|B_{ik}(x)|$ vanishes and which do not fall under any of the cases treated by the authors just referred to. In this category are problems of the type (1) which arise in connection with the calculus of variations, all of which have the property of self-adjointness described above. The importance of this class is evident when one considers the fact that the boundary problems of mathematical physics almost invariably belong to it. Another case is the system of the type (1) equivalent to the well known boundary value problem for a single linear differential equation of the nth order. This system does not fall under the theories presented in the papers mentioned above, and only self-adjoint equations of this sort give rise to self-adjoint problems of the type considered in the following pages. The interrelationships just mentioned between problems of various types are not elucidated in detail in this paper. I expect to return to them on another occasion.

The properties of the system (1) which justify expansion theorems analogous to those well known for other problems have not so far been clearly classified and analyzed. They seem to depend primarily upon the character of the matrix of functions $B_{ik}(x)$. A further study of this question, and a correlation of the methods and results hitherto attained, is desirable.

The methods used in the following pages are developed independently of the theory of linear integral equations though there are many points of contact. In Section 4 below it is shown that the equations (1) are equivalent to a certain system of linear integral equations for the functions $y_i(x)$, a result which is well known. The kernel matrix of these integral equations is not in general symmetric. When (1) is self-adjoint, however, every solution

$y_i(x)$ of the integral equations goes into a solution of the "associated" system of integral equations by means of a transformation of the form

$$u_i = \sum_{\alpha=1}^{n} S_{i\alpha} y_\alpha = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} B_{\beta i} T_{\beta\alpha} y_\alpha$$

for which the matrix $\|S_{ik}\|$ is symmetric. If the determinant $|B_{ik}|$ is different from zero the same is true of $|S_{ik}|$, and the system of integral equations equivalent to (1) can be reduced to one with a symmetric kernel matrix by means of a suitable transformation. The theory of the boundary value problem is then a corollary to the theory of a system of integral equations with symmetric kernel matrix such as has been developed, for example, by E. H. Moore in his "general analysis." The case when the determinant $|B_{ik}|$ vanishes includes some of the most important boundary value problems, as has been explained in the preceding paragraphs. For some of these it is possible to reduce the number of functions $y_i(x)$ by transformation and thereby to change the problem into an equivalent one whose corresponding integral equations have a symmetric kernel matrix, but it does not seem possible always to make such a transformation.

1. **Adjoint systems.** In the following pages the notations of tensor analysis will be used. It is understood that the indices $i, j, k, \cdots, \alpha, \beta, \gamma, \cdots$ have the range $1, 2, \cdots, n$ and that an expression such as $A_{i\alpha}B_{k\alpha}$ stands for the sum

$$A_{i\alpha}B_{k\alpha} = A_{i1}B_{k1} + \cdots + A_{in}B_{kn}$$

taken with respect to the repeated index α. The index α may be called an umbral index.* Such indices will usually be denoted by Greek letters.

As a matter of formal algebra, consider the four linear expressions

$$\begin{aligned} s_i(y) &= M_{i\alpha}y_\alpha(a) + N_{i\alpha}y_\alpha(b), & t_i(z) &= P_{\alpha i}z_\alpha(a) + Q_{\alpha i}z_\alpha(b), \\ \bar{s}_i(y) &= \bar{M}_{i\alpha}y_\alpha(a) + \bar{N}_{i\alpha}y_\alpha(b), & \bar{t}_i(z) &= \bar{P}_{\alpha i}z_\alpha(a) + \bar{Q}_{\alpha i}z_\alpha(b), \end{aligned} \tag{4}$$

formed for two sets of functions $y_i(x)$, $z_i(x)$. The coefficients in these expressions are real constants and the symbols on the left are merely notations for the sums on the right. If the coefficients in the sums $s_i(y)$ are given, with matrix of rank n, then a matrix of coefficients for $t_i(z)$ can always be selected also of rank n and so that the relations

$$M_{i\alpha}P_{\alpha k} - N_{i\alpha}Q_{\alpha k} = 0 \tag{5}$$

---

* Murnaghan, *Vector Analysis and the Theory of Relativity*, p. 7.

are satisfied. The coefficients of the auxiliary sums $\bar{s}_i(y)$ and $\bar{t}_i(z)$ can then be chosen so that the two matrices

$$\left\|\begin{matrix} M_{ik} & N_{ik} \\ \bar{M}_{ik} & \bar{N}_{ik} \end{matrix}\right\| \qquad \left\|\begin{matrix} -\bar{P}_{ik} & -P_{ik} \\ \bar{Q}_{ik} & Q_{ik} \end{matrix}\right\|, \tag{6}$$

with i constant in the rows and k in the columns, are reciprocals. From the relationships between the coefficients of these matrices found by multiplying them in reverse order it follows readily that the equation

$$s_\alpha(y)\bar{t}_\alpha(z) + \bar{s}_\alpha(y)t_\alpha(z) = y_\alpha(x)z_\alpha(x)\Big|_a^b \tag{7}$$

is an identity for all systems $y_i(x)$, $z_i(x)$. This is an equation which will be frequently useful. When the coefficients of the sums $s_i(y)$ are given the coefficients of the $t_i(z)$, chosen as above indicated, form $2n$ linearly independent solutions u_i, v_i of the equations

$$M_{i\alpha}u_\alpha - N_{i\alpha}v_\alpha = 0.$$

All other solutions of these equations are expressible linearly in terms of these n, and it follows readily that two different choices of the coefficients in $t_i(z)$ give systems of equations $t_i(z) = 0$ which are equivalent.

The systems of differential equations and boundary conditions considered in this paper may be written in the form

$$y_i'(x) = A_{i\alpha}(x)y_\alpha(x), \quad s_i(y) = M_{i\alpha}y_\alpha(a) + N_{i\alpha}y_\alpha(b) = 0. \tag{8}$$

The coefficients $A_{ik}(x)$ are supposed to be real single-valued continuous functions on the interval $a \leqq x \leqq b$; the matrix $\|M_{ik}, N_{ik}\|$ of coefficients in $s_i(y)$ is of rank n; and solutions $y_i(x)$ of the equations, continuous with their derivatives on the interval ab, are sought. A system adjoint to (8) is by definition one of the form

$$z_i'(x) = -A_{\alpha i}(x)z_\alpha(x), \quad t_i(z) = P_{\alpha i}z_\alpha(a) + Q_{\alpha i}z_\alpha(b) = 0 \tag{9}$$

where $t_i(z)$ has been formed from $s_i(y)$ in the manner described in the preceding paragraph.*

Let $\|Y_{ik}(x)\|$ be a matrix of functions whose columns form n linearly independent solutions of the differential equations in the system (8). The most general solution of these equations has the form $y_i(x) = Y_{i\alpha}c_\alpha$ where the coefficients c_i are constants. The solution $Y_{1i}, \cdots, Y_{ni}$ will be denoted simply by Y_i when no confusion results, and the determinant whose elements

* This definition is that of Bounitzky, p. 73, and is analogous to that of Birkhoff for linear equations of the nth order, these Transactions, vol. 9 (1908), p. 375.

are $D_{ik}=s_i(Y_k)$ will be denoted by D. For the adjoint system (9) the corresponding notations will be Z_{ik}, E. The first two theorems to be proved below are analogous to similar theorems for linear integral equations and for other types of boundary value problems, and are already in the literature.*

THEOREM 1. *If the determinant D is different from zero the only solution of the boundary value problem* (8) *is $y_i(x)=0$. If D has rank $n-r$ then the problem has r and only r linearly independent non-vanishing solutions.*

This is easy to see, since every solution of the boundary value problem must have the form $y_i=Y_{i\alpha}c_\alpha$ with coefficients c_i satisfying the linear equations $D_{i\alpha}c_\alpha=s_i(Y_\alpha)c_\alpha=0$.

THEOREM 2. *The number of linearly independent non-vanishing solutions of the original system* (8) *is always the same as the number for the adjoint system* (9).

To prove this one may note first that for every pair $y_i(x)$, $z_i(x)$ of solutions of the differential equations in (8) and (9) the equations

$$\begin{aligned} z_\beta y_\beta' + z_\beta' y_\beta &= A_{\beta\alpha}y_\alpha z_\beta - A_{\alpha\beta}y_\beta z_\alpha = 0, \\ y_\alpha(x)z_\alpha(x)\Big|_a^b &= 0 \end{aligned} \tag{10}$$

hold true. Suppose then that $U_{1p}, \cdots, U_{np}$ $(p=1, 2, \cdots, r)$ are r linearly independent solutions of the boundary value problem (8). The r sets $\bar{s}_i(U_p)$ $(p=1, 2, \cdots, r)$ are also linearly independent, since otherwise there would be a set of functions $y_i=U_{ip}c_p$, with constants c_p not all zero, making the expressions $\bar{s}_i(y)$, as well as the sums $s_i(y)$, all zero. Since the determinant of the first matrix (6) is not zero the equations (4) would then imply $y_i(a)=y_i(b)=0$, and the solution $y_i(x)$ would have to vanish identically, which is impossible when the solutions U_p are linearly independent. From the relations (7) and (10) it follows now that

$$\bar{s}_\alpha(U_p)t_\alpha(Z_k) = 0 \qquad (k = 1,2, \cdots, n;\ p = 1,2, \cdots, r).$$

Hence the determinant E whose elements are $E_{ik}=t_i(Z_k)$ has at most rank $n-r$. By similar reasoning it follows that when E has rank $n-r$ the rank of D is at most $n-r$, so that D and E have the same rank, and the number of linearly independent solutions of the systems (8) and (9) is the same.

THEOREM 3. *If the determinant D is different from zero the non-homogeneous system*

$$y_i'(x) = A_{i\alpha}(x)y_\alpha(x) + g_i(x), \qquad s_i(y) = h_i, \tag{11}$$

* Bounitzky, p. 77; Birkhoff-Langer, p. 64–5; Hurwitz, p. 526; Camp, p. 30.

where the functions $g_i(x)$ are continuous on the interval ab, has one and only one solution.

The differential equations of the system (11) have as their general solution

$$y_i(x) = Y_{i0}(x) + Y_{i\alpha}(x)c_\alpha \tag{12}$$

where Y_0 is a particular solution and the sets Y_α are as before linearly independent solutions of the equations with $g_i(x)=0$. Since $D=|s_i(Y_k)|$ is different from zero the constants c_i can be chosen in one and but one way so that

$$s_i(y) = s_i(Y_0) + s_i(Y_\alpha)c_\alpha = h_i.$$

THEOREM 4. *If the determinant D has rank $n-r$ then the system* (11) *has solutions if and only if the equation*

$$\int_a^b z_\alpha(x)g_\alpha(x)dx = 0 \tag{13}$$

is satisfied for every solution $z_i(x)$ of the adjoint system (9). *The most general solution of* (11) *is then*

$$y_i(x) = y_i^*(x) + c_1U_{i1}(x) + \cdots + c_rU_{ir}(x)$$

where $y_i^*(x)$ is a particular solution and the sets U_{ip} $(p=1,\cdots,r)$ are r linearly independent solutions of the original system (8).

If $y_i(x)$ and $z_i(x)$ are solutions of the differential equations in the systems (11) and (9), then one readily verifies that

$$y_\alpha' z_\alpha + y_\alpha z_\alpha' = g_\alpha z_\alpha, \tag{14}$$

and hence, with the help of equations (7), that

$$\int_a^b g_\alpha z_\alpha dx = y_\alpha z_\alpha \Big|_a^b = s_\alpha(y)\bar{t}_\alpha(z) + \bar{s}_\alpha(y)t_\alpha(z). \tag{15}$$

If y and z satisfy the boundary conditions in (11) and (9) the value of this expression is zero.

Suppose, conversely, that equation (13) holds true for every solution $z_i(x)$ of the adjoint system (9). Every solution of the original system (11) must have the form (12) with constants c_i satisfying the equations

$$s_i(y) = s_i(Y_0) + c_\alpha s_i(Y_\alpha) = 0.$$

By the argument of the proof of Theorem 2 it follows that the sets $\bar{t}_i(z)$, formed for r linearly independent solutions $z_i(x)$ of the adjoint system (9),

are themselves r linearly independent solutions of the linear equations whose coefficients are the columns of the determinant $D=|s_i(Y_k)|$. Hence the last equations have solutions c_i if and only if the conditions $\bar{t}_\alpha(z)\ s_\alpha(Y_0)=0$ are satisfied for every solution $z_i(x)$ of the adjoint system (9). But from equation (15) with y replaced by Y_0, and from the boundary conditions $t_i(z)=0$, it follows then that solutions c_i will certainly exist since the conditions

$$0 = \int_a^b g_\alpha z_\alpha dx = s_\alpha(Y_0)t_\alpha(z)$$

are satisfied.

The last statement of the theorem is true since if $y_i(x)$ and $y_i^*(x)$ are both solutions of the non-homogeneous system (11), their difference $y_i(x) - y_i^*(x)$ satisfies the original homogeneous system (8).

2. **Self-adjoint systems.** Consider now a system of the form

$$(16) \qquad y_i' = (A_{i\alpha} + \lambda B_{i\alpha})y_\alpha, \qquad s_i(y) = M_{i\alpha}y_\alpha(a) + N_{i\alpha}y_\alpha(b) = 0$$

containing a parameter λ linearly and having its coefficients $A_{ik}(x)$, $B_{ik}(x)$ all continuous on the interval $a \leqq x \leqq b$. The functions $B_{ik}(x)$ are by hypothesis not all identically zero. An adjoint system has the form

$$(17) \qquad z_i' = -(A_{\alpha i} + \lambda B_{\alpha i})z_\alpha, \qquad t_i(z) = P_{\alpha i}z_\alpha(a) + Q_{\alpha i}z_\alpha(b) = 0,$$

where $s_i(y)$ and $t_i(z)$ are related as in the preceding section.

The existence theorems for differential equations tell us that there exists a matrix $\|Y_{ik}(x, \lambda)\|$ whose columns are n linearly independent solutions of the differential equations in the system (16) and whose elements with their derivatives with respect to x are expressible as permanently convergent power series in λ. The determinant $D(\lambda)$ formed for this system is therefore also representable by a permanently convergent power series. The roots of $D(\lambda)$ are called the *characteristic values* of the parameter λ, and the non-vanishing solutions $y_i(x)$ of the system (16) corresponding to such values are called *characteristic solutions*. The corresponding notations for the adjoint system (17) are $Z_{ik}(x, \lambda)$ and $E(\lambda)$. It is well known that the roots of a permanently convergent power series $D(\lambda)$ are finite or at most denumerably infinite in number. The following theorem is an immediate consequence of Theorems 1 and 2 of the preceding section.

THEOREM 5. *The characteristic parameter values for the boundary value problem* (16) *are identical with those for the adjoint problem* (17). *The number of linearly independent characteristic solutions corresponding to a particular characteristic value* λ *is the same for both problems.*

THEOREM 6. *If* $\lambda_1 \neq \lambda_2$ *are characteristic values and* $y_i(x)$, $z_i(x)$ *corresponding characteristic solutions of* (16) *and* (17), *then*

$$\int_a^b B_{\alpha\beta}(x) y_\alpha(x) z_\beta(x) dx = 0. \tag{18}$$

For from the differential equations in (16) and (17) with λ_1 and λ_2 substituted it follows that

$$y_\alpha' z_\alpha + y_\alpha z_\alpha' = \lambda_1 B_{\alpha\beta} y_\beta z_\alpha - \lambda_2 B_{\beta\alpha} y_\alpha z_\beta,$$

and hence, from the boundary conditions $s_i(y) = t_i(z) = 0$ and equation (7), that

$$0 = y_\alpha z_\beta \Big|_a^b = (\lambda_1 - \lambda_2) \int_a^b B_{\beta\alpha} y_\alpha z_\beta dx.$$

In the following pages a set of relations of the form $z_i = T_{i\alpha}(x) y_\alpha$ will be called a transformation if the functions $T_{ik}(x)$ are real, single-valued, and have continuous derivatives on the interval ab, and if the determinant $|T_{ik}(x)|$ is different from zero on that interval. The coefficients of the inverse transformation will be denoted by $\overset{-1}{T}_{ik}(x)$.

Definition of a self-adjoint system. The boundary value problem (16) is said to be *self-adjoint* if the differential equations and also the boundary conditions of its adjoint (17) are equivalent to its own for all values of λ by means of a transformation $z_i = T_{i\alpha}(x) y_\alpha$.

THEOREM 7. *In order that the problem* (16) *shall be self-adjoint it is necessary and sufficient that there shall exist a transformation* $T_{ik}(x)$ *such that*

$$T_{i\alpha} A_{\alpha k} + A_{\alpha i} T_{\alpha k} + T_{ik}' \equiv 0, \qquad T_{i\alpha} B_{\alpha k} + B_{\alpha i} T_{\alpha k} \equiv 0, \tag{19}$$

$$M_{i\alpha} \overset{-1}{T}_{\alpha\beta}(a) M_{k\beta} = N_{i\alpha} \overset{-1}{T}_{\alpha\beta}(b) N_{k\beta}. \tag{20}$$

To prove the first two of these relations one can verify readily that the transformation $z_i = T_{i\alpha} y_\alpha$ takes the differential equations of the system (16) into the set

$$z_i' = [T_{i\alpha} A_{\alpha\beta} \overset{-1}{T}_{\beta\gamma} + \lambda T_{i\alpha} B_{\alpha\beta} \overset{-1}{T}_{\beta\gamma} + T_{i\beta}' \overset{-1}{T}_{\beta\gamma}] z_\gamma.$$

If these are equivalent to the differential equations in the adjoint system (17) for all values of λ, the equations (19) follow at once.

By the transformation $z_i = T_{i\alpha} y_\alpha$ the boundary conditions of the adjoint problem take the form

$$P_{\alpha i} T_{\alpha\beta}(a) y_\beta(a) + Q_{\alpha i} T_{\alpha\beta}(b) y_\beta(b) = 0.$$

For these to be equivalent to the original boundary conditions $s_i(y)=0$ it is necessary and sufficient that a matrix of constants C_{ik} with determinant different from zero exists such that

$$(21) \qquad P_{\alpha i}T_{\alpha k}(a) = C_{i\alpha}M_{\alpha k}, \qquad Q_{\alpha i}T_{\alpha k}(b) = C_{i\alpha}N_{\alpha k}.$$

With the help of the relations (5) it follows that the preceding equations imply

$$C_{i\alpha}[M_{\alpha\beta}\overset{-1}{T}_{\beta\gamma}(a)M_{k\gamma} - N_{\alpha\beta}\overset{-1}{T}_{\beta\gamma}(b)N_{k\gamma}] = 0.$$

Since $|C_{ik}|\neq 0$ the equations (20) of the theorem follow at once. Conversely, since the matrix $\|M_{ik}, N_{ik}\|$ is of rank n, the equations (5) and (20) imply relations of the form (21), so that the boundary conditions of the two systems are surely equivalent.

THEOREM 8. *For a self-adjoint system* (16) *two characteristic solutions* $y_i(x), \bar{y}_i(x)$ *corresponding to distinct characteristic valus* $\lambda, \bar{\lambda}$ *satisfy the equation*

$$(22) \qquad \int_a^b S_{\alpha\beta}(x)y_\alpha(x)\bar{y}_\beta(x)dx = 0$$

where $S_{ik}=T_{\alpha i}\, B_{\alpha k}$.

This is an immediate consequence of the equation (18) and the transformation $z_i=T_{i\alpha}y_\alpha$.

For the following definition it is important to note that when the matrix $\|S_{ik}\|$ is symmetric the bilinear form $S_{\alpha\beta}f_\alpha\bar{f}_\beta$, formed for a set of numbers f_i and their conjugate imaginaries $\bar{f}_i$, is always real, since such a form is identical with its conjugate. The functions $g_\alpha(x)$ in the definition are supposed to be continuous on the interval ab.

Definition of a definitely self-adjoint boundary value problem. A problem (16) is said to be *definitely self-adjoint* if the matrix $\|S_{ik}(x)\|$ is symmetric and the bilinear form $S_{\alpha\beta}(x)\, f_\alpha\bar{f}_\beta$ is positive or zero at every point of the interval ab, and if furthermore this form vanishes identically for a set of solutions $f_i(x)$ of a system of equations of the type

$$(23) \qquad f_i'(x) = A_{i\alpha}(x)f_\alpha(x) + B_{i\alpha}(x)g_\alpha(x)$$

only when the functions $f_i(x)$ are all identically zero.

The conditions of the definition will surely be satisfied if the quadratic form $S_{\alpha\beta}(x)\, f_\alpha f_\beta$ is positive definite at every point of the interval ab, since the bilinear form $S_{\alpha\beta}\, f_\alpha\bar{f}_\beta$ is then always positive for non-vanishing arguments f_i, as one readily verifies. In that case the determinant $|S_{ik}|=|T_{ik}|\ |B_{ik}|$ is

everywhere different from zero. The same is therefore true of $|B_{ik}|$, and the second equation (19) shows that

$$0 = S_{ik} - S_{ki} = T_{\alpha i}B_{\alpha k} - T_{\alpha k}B_{\alpha i} = (T_{\alpha i} + T_{i\alpha})B_{\alpha k}.$$

It follows readily that the matrix $\|T_{ik}\|$ is skew-symmetric. Since a skew-symmetric determinant of odd order always vanishes, and since the determinant $|T_{ik}|$ must be different from zero, it is clear that this case can arise only when n is even. It should be noted, however, that the definition is applicable to cases when the determinant $|B_{ik}|$ vanishes and the quadratic form $S_{\alpha\beta}f_\alpha f_\beta$ is not definite. Important special cases of this sort are the boundary value problems arising from the calculus of variations and the problem which arises when a boundary value problem for a self-adjoint linear differential equation of the nth order, of a type hitherto often studied,* is transformed into one of the type (16).

If the bilinear form $S_{\alpha\beta}f_\alpha\bar{f}_\beta$ is non-positive it can always be replaced by one which is non-negative by using the transformation with coefficients $-T_{ik}$ instead of T_{ik}. The requirement of symmetry for the matrix $\|S_{ik}\|$ is also not as stringent as it perhaps appears to be at first sight. If the equations (19) have a system of solutions T_{ik} then the systems $\bar{T}_{ik}=T_{ki}$ and $\bar{T}_{ik}-T_{ik}$ are also solutions. The matrix of elements $\bar{T}_{ik}-T_{ik}=T_{ki}-T_{ik}$ is skew-symmetric, and one readily verifies by means of the relations (19) that for a skew-symmetric system T_{ik} the elements S_{ik} have the symmetry required. If the matrix of elements $T_{ki}-T_{ik}$ is to be useful for a transformation, however, the determinant $|T_{ki}-T_{ik}|$ must be different from zero.

When the problem (16) is definitely self-adjoint the elements S_{ik} are expressible in the various forms

$$S_{ik} = T_{\alpha i}B_{\alpha k} = T_{\alpha k}B_{\alpha i} = - T_{i\alpha}B_{\alpha k} \tag{24}$$

as one readily verifies from the symmetry of S_{ik} and the relations (19).

3. **Properties of self-adjoint systems.** Boundary value problems of the type (16) which are definitely self-adjoint have many properties analogous to those of linear integral equations with symmetric kernel functions, as indicated in the following theorems.

THEOREM 9. *For a definitely self-adjoint boundary value problem* (16) *all roots of the characteristic determinant* $D(\lambda)$ *are real and the linearly independent characteristic solutions corresponding to each root may be chosen real.*

---

* See, for example, Darboux, *Théorie des Surfaces*, vol. 2, p. 109; Birkhoff, loc. cit., p. 373; Bounitzky, p. 88.

For suppose λ a root of $D(\lambda)$ and $y_i(x)$ a non-vanishing solution of the system (16) corresponding to it, and let $\bar{\lambda}$ and $\bar{y}_i(x)$ be their conjugate imaginaries. If λ were not real, equation (22) of Theorem 8 would require the bilinear form $S_{\alpha\beta}y_\alpha\bar{y}_\beta$ to vanish identically in x, which is impossible when the solution $y_i(x)$ is not identically zero. Hence the root λ is real. But if λ is real then the real and imaginary parts of $y_i(x)$ are separately solutions of the system (16), and it is evident that a linearly independent set of real characteristic solutions corresponding to λ can be selected.

THEOREM 10. *The index of each characteristic number λ_0, i.e., the number of linearly independent characteristic solutions $y_i(x)$ corresponding to it, is equal to the multiplicity of λ_0 as a root of $D(\lambda)$.*

Suppose that $D(\lambda)=|s_i[Y_k(x,\lambda)]|$ has rank $n-r$ at a particular value λ_0. By replacing the solutions $Y_k(x,\lambda)$ by suitably selected linear combinations of them with constant coefficients it may be brought about that for $\lambda=\lambda_0$ the expressions $s_i(Y_p)$ $(p=1,\cdots,r)$ all vanish, while the matrix of elements $s_i(Y_q)$ $(q=r+1,\cdots,n)$ has rank $n-r$. All derivatives of $D(\lambda)$ of order less than r will then clearly vanish at $\lambda=\lambda_0$, and the rth will have the value

$$D^{(r)}(\lambda_0)=|\,s_i(Y_{1\lambda}),\cdots,s_i(Y_{r\lambda}),s_i(Y_{r+1}),\cdots,s_i(Y_n)\,| \tag{25}$$

where the subscript λ indicates derivatives. If this expression vanished there would be a linear combination

$$y_i=(c_1Y_{i1\lambda}+\cdots+c_rY_{ir\lambda})+(c_{r+1}Y_{i,r+1}+\cdots+c_nY_{in})$$

for which all the numbers $s_i(y)$ would vanish at $\lambda=\lambda_0$. The constants $c_1,\cdots,c_r$ could not all be zero because the rank of the last $n-r$ columns of $D^{(r)}(\lambda_0)$ is $n-r$. The functions y_{i1} whose derivatives for λ are in the first parenthesis would therefore not vanish identically. For $\lambda=\lambda_0$ they would satisfy the system (16) and also the equations

$$y'_{i1\lambda}=(A_{i\alpha}+\lambda B_{i\alpha})y_{\alpha1\lambda}+B_{i\alpha}y_{\alpha1}.$$

The set y_{i2} defined by the second parenthesis would satisfy the differential equations of the system (16), and it follows readily that the functions y_i themselves would for $\lambda=\lambda_0$ be solutions of the non-homogeneous system

$$y_i'=(A_{i\alpha}+\lambda_0B_{i\alpha})y_\alpha+B_{i\alpha}y_{\alpha1},\qquad s_i(y)=0.$$

The functions $z_{i1}=T_{i\alpha}y_{\alpha1}$ would satisfy the adjoint equations (17), and from equations (13) of Theorem 4 it would follow that

$$\int_a^b z_{\alpha1}B_{\alpha\beta}y_{\beta1}dx=\int_a^b y_{\alpha1}S_{\alpha\beta}y_{\beta1}dx=0.$$

This could not be true, however, since the functions y_{i1} would not all vanish, as was seen above. It follows therefore that the derivative (25) is different from zero and that λ_0 has its multiplicity equal to its index.

THEOREM 11. *For a system of functions $f_i(x)$ continuous on the interval ab and satisfying the condition*

$$\int_a^b y_\alpha(x)S_{\alpha\beta}(x)f_\beta(x)dx = 0 \tag{26}$$

with every characteristic solution $y_i(x)$ of the boundary value problem (16), *the functions $B_{i\alpha}(x)f_\alpha(x)$ all vanish identically.**

To prove this let $f_i(x)$ be a set with the properties described in the theorem. According to Theorems 3 and 4 the non-homogeneous system

$$y_i' = (A_{i\alpha} + \lambda B_{i\alpha})y_\alpha + B_{i\alpha}f_\alpha, \qquad s_i(y) = 0 \tag{27}$$

then has solutions for every value of λ, since the equations (26) imply the conditions analogous for this case to equations (13) of Theorem 4. When $D(\lambda)\neq 0$ there is a unique solution and one verifies readily that it consists of the functions

$$y_i(x,\lambda) = \frac{1}{D(\lambda)}\begin{vmatrix} Y_{i0} & Y_{i1} & \cdots & Y_{in} \\ s_1(Y_0) & s_1(Y_1) & \cdots & s_1(Y_n) \\ \cdot & \cdot & \cdots & \cdot \\ s_n(Y_0) & s_n(Y_1) & \cdots & s_n(Y_n) \end{vmatrix} \tag{28}$$

where $Y_{i0}(x, \lambda)$ is a particular solution of the differential equations in the system (27).

Near a root λ_0 of multiplicity r of $D(\lambda)$ the functions $y_i(x, \lambda)$ are still well-defined and analytic in λ. For, in the first place, one can add constant multiples of the last n columns of the determinant to the first column in such a way that the resulting functions Y_{i0} satisfy the conditions $s_i(Y_0)=0$ at $\lambda=\lambda_0$. In the second place the r linearly independent solutions $z_i(x)$ of the adjoint system (17) for $\lambda=\lambda_0$ provide r linearly independent sets $\bar{t}_\alpha(z)$, as in the proof of Theorem 2, such that

$$s_\alpha(Y_1)\bar{t}_\alpha(z) = \cdots = s_\alpha(Y_n)\bar{t}_\alpha(z) = 0.$$

One can therefore replace r of the rows of the determinant in the expressions (28) by r linear combinations of its n rows which vanish at $\lambda=\lambda_0$.

* The idea underlying the proof of this theorem is very well known. It has been used, for example, by E. Schmidt, p. 18; Hurwitz, p. 539; and Camp, p. 37.

It is clear that the determinant in (28) has the same factor $(\lambda-\lambda_0)^r$ as $D(\lambda)$, and hence that the functions $y_i(x, \lambda)$ are analytic near λ_0 as well as near the values of λ at which $D(\lambda)\neq 0$.

The functions $y_i(x, \lambda)$ are representable by permanently convergent power series in λ of the form

$$y_i(x,\lambda) = u_{i0}(x) + u_{i1}(x)\lambda + u_{i2}(x)\lambda^2 + \cdots. \tag{29}$$

By substituting these series in the equations (27) and comparing coefficients of λ it is found that the coefficients $u_{i\mu}$ $(\mu=0, 1, 2, \cdots)$ satisfy the equations

$$u'_{i\mu} = A_{i\alpha}u_{\alpha\mu} + B_{i\alpha}u_{\alpha,\mu-1}, \qquad s_i(u_\mu) = 0 \tag{30}$$

in which it is agreed that $u_{i,-1}=f_i$. With the help of equations (30) and (19) one verifies also that the functions $v_{i\nu} = T_{i\alpha}u_{\alpha\nu}$ are solutions of the system

$$v'_{i\nu} = -A_{\alpha i}v_{\alpha\nu} - B_{\alpha i}v_{\alpha,\nu-1}, \qquad t_i(v_\nu) = 0,$$

where $v_{i,-1}=T_{i\alpha}f_\alpha$, and it follows that

$$u'_{\alpha\mu}v_{\alpha\nu} + u_{\alpha\mu}v'_{\alpha\nu} = u_{\beta,\mu-1}B_{\alpha\beta}v_{\alpha\nu} - u_{\alpha\mu}B_{\beta\alpha}v_{\beta,\nu-1}.$$

This result and the equations (7) and (24) now justify the relation

$$0 = u_{\alpha\mu}v_{\alpha\nu}\Big|_a^b = \int_a^b u_{\alpha,\mu-1}S_{\alpha\beta}u_{\beta\nu}dx - \int_a^b u_{\alpha\mu}S_{\alpha\beta}u_{\beta,\nu-1}dx.$$

The analogue of a well known inequality of Schwarz

$$\left(\int_a^b u_{\alpha,\mu-1}S_{\alpha\beta}u_{\alpha,\mu+1}dx\right)^2 \leqq \int_a^b u_{\alpha,\mu-1}S_{\alpha\beta}u_{\alpha,\mu-1}dx\int_a^b u_{\alpha,\mu+1}S_{\alpha\beta}u_{\alpha,\mu+1}dx,$$

which will be proved in the next section, together with the last equation, show that the constants

$$W_\mu = \int_a^b u_{\alpha0}S_{\alpha\beta}u_{\alpha\mu}dx$$

have the properties

$$W_{\mu+\nu} = \int_a^b u_{\alpha\mu}S_{\alpha\beta}u_{\alpha\nu}dx, \qquad W_{2\mu-2}W_{2\mu+2} \geqq W_{2\mu}^2. \tag{31}$$

The series

$$W_0 + W_1\lambda + \cdots, \qquad W_0 + W_2\lambda^2 + \cdots, \tag{32}$$

the first of which is found by integrating the expression $u_{\alpha0}S_{\alpha\beta}y_\beta$ found from (29), converge for every value of λ. Some coefficient in the second series must vanish. Otherwise it would follow from the inequality (31) that

$W_{2\mu} \geqq W_0(W_2/W_0)^\mu$, and the second series (32) would not converge for $\lambda = (W_2/W_0)^{\frac{1}{2}}$. But if a particular coefficient $W_{2\mu}$ vanishes the inequalities (31) imply that all the preceding ones of even index are also zero. From the equation $W_0 = 0$ it follows, however, that the functions u_{i0} all vanish identically, and the first of the equations (30) shows that the functions $u_{i,-1} = B_{i\alpha}f_\alpha$ do the same.

COROLLARY 1. *If the determinant* $|B_{ik}(x)|$ *is different from zero on the interval ab, as in the case when the matrix* $\|S_{ik}\| = \|T_{\alpha i}\, B_{\alpha k}\|$ *is positive definite at each point of ab, then the set of functions* $f_i(x) = 0$ *is the only one satisfying the relation* (26) *with all characteristic solutions* $y_i(x)$ *of the boundary value problem* (16).

COROLLARY 2. *The only set of solutions* $f_i(x)$ *of a system of equations of the form*

$$f_i' = A_{i\alpha}f_\alpha + B_{i\alpha}g_\alpha \tag{33}$$

satisfying the relations (26) *with all characteristic solutions* $y_i(x)$ *is the set* $f_i(x) \equiv 0$.

With the help of the last corollary it is possible to prove

THEOREM 12. *The totality of characteristic solutions and characteristic constants of the boundary value problem* (16) *are denumerably infinite in number and may be represented by the symbols* $y_{i\nu}(x), \lambda_\nu (\nu = 1, 2, \cdots)$. *These functions may furthermore be chosen normed and orthogonal in the sense that*

$$\int_a^b y_{\alpha\mu}S_{\alpha\beta}y_{\beta\nu}dx = \delta_{\mu\nu} \qquad (\delta_{\mu\mu} = 1,\ \delta_{\mu\nu} = 0 \text{ if } \mu \neq \nu). \tag{34}$$

Let the functions g_α in equations (33) have the form $g_i = d_\rho\, g_{i\rho}$ $(\rho = 1, \cdots, p)$. The functions $g_{i\rho}$ can be chosen so that the p systems $B_{i\alpha}\, g_{\alpha\rho}$ $(\rho = 1, \cdots, p)$ are linearly independent, since the functions B_{ik} do not all vanish identically. The system (33) has a solution of the form $f_i = d_\rho f_{i\rho}$ where the system $f_{i\rho}$ for p fixed is a particular solution of the equations (33) with g_i replaced by $g_{i\rho}$. The solutions $f_{i\rho}$ are readily seen to be linearly independent since the sets $B_{i\alpha}\, g_{\alpha\rho}(\rho = 1, \cdots, p)$ have this property. If there were a finite number only of characteristic solutions of the boundary value problem (16) then for a sufficiently large value of p, constants d_ρ not all zero could always be selected so that the functions $f_i = d_\rho f_{i\rho}$ would satisfy the relations (26) with every such solution, and this would contradict Corollary 2 above. Hence there must be an infinity of characteristic solutions, and the infinity is denumerable since the roots of the permanently convergent series $D(\lambda)$ are denumerable.

Since no one of the sets $y_{i\mu}(x)$ for fixed μ is identically zero the integrals (26) for $\nu=\mu$ are all different from zero. The characteristic functions $y_{i\mu}(x)$ may therefore be normed and orthogonalized as described in the theorem by a well known process.*

4. Expansion theorems. With the help of the theorems of the last section it is possible to deduce some very general expansion theorems for sets of functions $f_i(x)$. The characteristic solutions $y_{i\mu}(x)$ of the definitely self-adjoint boundary value problem (16), appearing in these theorems, are supposed to be normed and orthogonal.

THEOREM 13. *For every solution $f_i(x)$ of a system of equations of the form*

$$f_i' = A_{i\alpha}f_\alpha + B_{i\alpha}g_\alpha, \qquad s_i(f) = 0, \tag{35}$$

in which the functions $g_i(x)$ are arbitrarily selected continuous functions on the interval ab, the series

$$\varphi_i(x) = \sum_{\nu=1}^{\infty} y_{i\nu}(x) \int_a^b y_{\alpha\nu}(\xi)S_{\alpha\beta}(\xi)f_\beta(\xi)d\xi \tag{36}$$

converge uniformly and the functions $B_{i\alpha}(f_\alpha-\varphi_\alpha)$ are all identically zero on ab.

The uniform convergence of the series will be proved in a later section. The rest of the theorem follows at once from Theorem 11 since the equations

$$\int_a^b y_{\alpha\mu}S_{\alpha\beta}(f_\beta - \varphi_\beta)dx = 0 \qquad (\mu = 1,2,\cdots)$$

are immediate consequences of the definition of the functions φ_i.

COROLLARY 1. *If the determinant $|B_{ik}(x)|$ is different from zero on the interval ab, as in the case when the matrix $\|S_{ik}\| = \|T_{\alpha i}B_{\alpha k}\|$ is positive definite at each point of ab, then for every set of functions $f_i(x)$ having continuous derivatives on this interval and satisfying the boundary conditions $s_i(f)=0$ the series* (36) *converge uniformly and represent the functions $f_i(x)$.*

In this case the equations (35) determine uniquely a set of functions $g_i(x)$ corresponding to the given functions $f_i(x)$, since the determinant $|B_{ik}|$ is different from zero.

COROLLARY 2. *If the functions $f_i(x)$ are solutions of a system* (35) *and the functions $g_i(x)$ solutions of a similar system*

$$g_i' = A_{i\alpha}g_\alpha + B_{i\alpha}h_\alpha, \qquad s_i(g) = 0, \tag{37}$$

* See, for example, E. Schmidt, p. 4.

in which the functions $h_i(x)$ *are continuous on ab, then the series* (36) *converge uniformly and represent the functions* $f_i(x)$ *on this interval.*

To prove this the equations (35), (17), and (7) can first be used to show that for every characteristic solution $y_{i\nu}(x)$ of the system (15) the corresponding functions $z_{i\nu} = T_{i\alpha} y_{\alpha\nu}$ satisfy the relations

$$z_{\alpha\nu} f_\alpha' + z_{\alpha\nu}' f_\alpha = z_{\alpha\nu} B_{\epsilon\beta} g_\beta - \lambda_\nu z_{\beta\nu} B_{\beta\alpha} f_\alpha,$$

$$0 = z_{\alpha\nu} f_\alpha \Big|_a^b = \int_a^b y_{\alpha\nu} S_{\alpha\beta} g_\beta dx - \lambda_\nu \int_a^b y_{\alpha\nu} S_{\alpha\beta} f_\beta dx. \tag{38}$$

Since the functions $g_i(x)$ satisfy equations (37) it will be seen that the series

$$\sum\nolimits_\nu \lambda_\nu y_{i\nu}(x) \int_a^b y_{\alpha\nu} S_{\alpha\beta} f_\beta d\xi = \sum\nolimits_\nu y_{i\nu}(x) \int_a^b y_{\alpha\nu} S_{\alpha\beta} g_\beta d\xi$$

converges uniformly by the same proof as that which shows the convergence of the series (36). It follows readily with the help of the differential equations (16) for the functions $y_{i\nu}(x)$ that the series of derivatives of the terms of (36) converges uniformly and represents $\varphi_i'(x)$. Further

$$\varphi_i'(x) = A_{i\alpha}\varphi_\alpha + B_{i\alpha} \sum\nolimits_\nu y_{\alpha\nu}(x) \int_a^b y_{\alpha\nu} S_{\alpha\beta} g_\beta d\xi. \tag{39}$$

Since the functions $B_{i\alpha}(f_\alpha - \varphi_\alpha)$ vanish identically, $(f_\alpha - \varphi_\alpha) S_{\alpha\beta}(f_\beta - \varphi_\beta)$ vanishes identically. Since from equations (35) and (39) the differences $f_\alpha - \varphi_\alpha$ satisfy an equation of the type (23) it follows that they are identically zero, which was to be proved.

5. **Green's matrix.** Consider again the system

$$y_i' = A_{i\alpha} y_\alpha, \quad s_i(y) = M_{i\alpha} y_\alpha(a) + N_{i\alpha} y_\alpha(b) = 0 \tag{40}$$

described in § 1, and suppose that its determinant D is different from zero.

Definition of the Green's matrix.* A matrix of functions $G_{ik}(x, \zeta)$, single-valued for $a \leq x \leq b$, $a \leq \xi \leq b$, and further such that they are continuous and have continuous first derivatives in x except at $x = \xi$, is called the *Green's matrix* of the system (16) if it has the properties

$$G_{ik}(\xi + 0, \xi) - G_{ik}(\xi - 0, \xi) = \delta_{ik} \quad (\delta_{ii} = 1, \delta_{ik} = 0 \text{ if } i \neq k), \tag{41}$$

$$\frac{\partial}{\partial x} G_{ik}(x, \xi) = A_{i\alpha}(x) G_{\alpha k}(x, \xi), \tag{42}$$

$$M_{i\alpha} G_{\alpha k}(a, \xi) + N_{i\alpha} G_{\alpha k}(b, \xi) = 0. \tag{43}$$

---

* The existence and uniqueness of the Green's matrix has been proved by several writers. See Bounitzky, p. 77; Birkhoff-Langer, pp. 66-70.

The properties (42) and (43) show that the columns of the matrix would be solutions of the system (16) of the kind demanded by the boundary value problem if it were not for their discontinuities at $x=\xi$. Let D_{ik} and Δ_{ik} be defined by the equations

$$(44)\qquad D_{ik} = M_{i\alpha}Y_{\alpha k}(a) + N_{i\alpha}Y_{\alpha k}(b), \quad \Delta_{ik} = M_{i\alpha}Y_{\alpha k}(a) - N_{i\alpha}Y_{\alpha k}(b).$$

The determinant $D=|D_{ik}|$ is different from zero by hypothesis. The functions

$$(45)\qquad G_{ik}(x,\xi) = \frac{1}{2}Y_{i\alpha}(x)\left[\frac{|x-\xi|}{x-\xi}\delta_{\alpha\gamma} + D^{-1}_{\alpha\beta}\Delta_{\beta\gamma}\right]Y^{-1}_{\gamma k}(\xi)$$

are then well-defined and evidently have the properties (41) and (42). Equation (43) follows readily with the help of the notations (44).

THEOREM 14. *The functions* $H_{ik}(x,\xi) = -G_{ki}(\xi,x)$ *defined by the equations* (45) *are the elements of the Green's matrix of the adjoint system* (17), *so that*

$$(46)\qquad G_{ki}(x,x-0) - G_{ki}(x,x+0) = \delta_{ki},$$

$$(47)\qquad \frac{\partial}{\partial\xi}G_{ki}(x,\xi) = -A_{\alpha i}(\xi)G_{k\alpha}(x,\xi),$$

$$(48)\qquad P_{\alpha i}G_{k\alpha}(x,a) + Q_{\alpha i}G_{k\alpha}(x,b) = 0.$$

By differentiating the equations $Y_{i\alpha}\,Y^{-1}_{\alpha k}=\delta_{ik}$ it can be shown that the rows of the matrix Y^{-1}_{ik} are solutions of the differential equations of the adjoint system (17), and the properties (46) and (47) are then evident from the form of the equations (45). The relations $M_{i\alpha}P_{\alpha k}-N_{i\alpha}Q_{\alpha k}=0$ between the elements of the reciprocal matrices (6) justify the equations

$$\Delta_{i\alpha}[P_{\beta k}Y^{-1}_{\alpha\beta}(a) + Q_{\alpha i}Y^{-1}_{\alpha\beta}(b)] = D_{i\alpha}[-P_{\beta k}Y^{-1}_{\alpha\beta}(a) + Q_{\beta k}Y^{-1}_{\alpha\beta}(b)]$$

and equation (48) can be proved with the help of them.

THEOREM 15. *If the functions* $g_i(x)$ *are continuous on the interval ab then every solution* $y_i(x)$ *of the non-homogeneous system*

$$(49)\qquad y_i' = A_{i\alpha}y_\alpha + g_i, \qquad s_i(y) = 0$$

is expressible in the form

$$(50)\qquad y_i(x) = \int_a^b G_{i\alpha}(x,\xi)g_\alpha(\xi)d\xi,$$

and conversely every system $y_i(x)$ *defined by equations* (50) *is a solution of* (49).

From the equations (49) for $x=\xi$, and equations (47), it follows that

$$\frac{\partial}{\partial\xi}[G_{k\alpha}(x,\xi)y_\alpha(\xi)] = G_{k\alpha}(x,\xi)g_\alpha(\xi).$$

Hence

$$G_{k\alpha}(x,x-0)y_\alpha(x) - G_{k\alpha}(x,a)y_\alpha(a) = \int_a^x G_{k\alpha}(x,\xi)g_\alpha(\xi)d\xi,$$

$$G_{k\alpha}(x,b)y_\alpha(b) - G_{k\alpha}(x,x+0)y_\alpha(x) = \int_x^b G_{k\alpha}(x,\xi)g_\alpha(\xi)d\xi.$$

With the help of the relations (46), (48), $s_i(y)=0$, and (7), the sum of these two equations gives equation (50).

The converse is easily proved by writing the expression (50) for $y_i(x)$ as the sum of integrals on the intervals ax and xb, differentiating with respect to x, and using the relations (46) and (42).

COROLLARY 1. *The Green's matrix is unique.*

For if there were two, say G_{ik} and L_{ik}, the equation

$$\int_a^b [G_{i\alpha}(x,\xi) - L_{i\alpha}(x,\xi)]g_\alpha(\xi)d\xi = 0$$

would be an identity in x for every set of functions $g_\alpha(\xi)$, and this can be true only if $G_{ik}=L_{ik}$.

COROLLARY 2. *When $\lambda=0$ is not a characteristic number then every solution $y_i(x)$ of the boundary value problem*

$$y_i' = (A_{i\alpha}+\lambda B_{i\alpha})y_i, \qquad s_i(y) = 0 \tag{51}$$

is a solution of the system of linear integral equations

$$y_i(x) = \lambda\int_a^b G_{i\alpha}(x,\xi)B_{\alpha\beta}(\xi)y_\beta(\xi)d\xi, \tag{52}$$

and conversely.

By replacing λ by $\lambda_1+\lambda$ in the equations (16) it can readily be brought about that the value $\lambda=0$ is not a characteristic number, if this is not already the case.

THEOREM 16. *If the boundary value problem* (16) *is self-adjoint then the system* (52) *is equivalent to the corresponding system*

$$z_i(x) = \lambda\int_a^b z_\alpha(\xi)B_{\alpha\beta}(\xi)G_{\beta i}(\xi,x)d\xi \tag{53}$$

for the adjoint problem by means of the transformation $z_i = T_{i\alpha}y_\alpha$. *The functions* G_{ik} *in this case satisfy the relations*

$$G_{\alpha i}(\xi,x)T_{\alpha k}(\xi) = -T_{i\alpha}(x)G_{\alpha k}(x,\xi). \tag{54}$$

Since the original boundary value problem and its adjoint are equivalent under the transformation $z_i = T_{i\alpha}y_\alpha$ it is evident that the systems (52) and (53) must have the same property.

To prove equation (54) one may first deduce the second of the systems

$$\begin{aligned} y_i' &= A_{i\alpha}y_\alpha + g_\alpha, & s_i(y) &= 0, \\ z_i' &= -A_{\alpha i}z_\alpha + T_{i\alpha}g_\alpha, & t_i(z) &= 0 \end{aligned}$$

from the first by means of the transformation $z_i = T_{i\alpha}y_\alpha$. From Theorem 15 it follows then that the systems

$$y_i(x) = \int_a^b G_{i\alpha}(x,\xi)g_\alpha(\xi)d\xi,$$

$$z_i(x) = \int_a^b H_{i\alpha}(x,\xi)T_{\alpha\beta}(\xi)g_\alpha(\xi)d\xi$$

are equivalent by this transformation. But from these one verifies that the equation

$$\int_a^b [H_{i\alpha}(x,\xi)T_{\alpha\beta}(\xi) - T_{i\alpha}(x)G_{\alpha\beta}(x,\xi)]g_\beta(\xi)d\xi = 0$$

must hold for all sets of functions g_α. Since $H_{ik}(x,\xi) = -G_{ki}(\xi,x)$ this proves equations (54).

If the elements of the kernel matrix of the system (52) are denoted by

$$K_{ij}(x,\xi) = G_{i\alpha}(x,\xi)B_{\alpha j}(\xi)$$

then this system and the one "associated" with it, according to the theory of integral equations, are

$$\begin{aligned} y_i(x) &= \lambda \int_a^b K_{i\alpha}(x,\xi)y_\alpha(\xi)d\xi, \\ u_i(x) &= \lambda \int_a^b K_{\alpha i}(\xi,x)u_\alpha(\xi)d\xi. \end{aligned} \tag{55}$$

When the boundary value problem (51) is self-adjoint the relations (54), with (24), show that

$$K_{\alpha i}(\xi,x)S_{\alpha k}(\xi) = S_{i\beta}(x)K_{\beta k}(x,\xi).$$

It is provable readily then that every solution $y_i(x)$ of the first of equations (55) goes into a solution $u_i(x)$ of the second by means of the transformation $u_i(x)=S_{i\beta}(x)\,y_\beta(x)$. If the determinant $|S_{ik}|$ is different from zero the equations (55) are completely equivalent by means of this transformation. In that case a symmetric matrix $U_{ik}(x)$ can be determined such that

$$U_{i\alpha}(x)U_{\alpha k}(x)=S_{ik}(x)^*$$

and the transformation $v_i=U_{i\alpha}y_\alpha$ takes the first system (55) into the system

$$v_i(x) = \lambda\int_a^b U_{i\alpha}(x)K_{\alpha\beta}(x,\xi)U^{-1}_{\beta\gamma}(\xi)v_\gamma(\xi)d\xi$$

whose kernel matrix is readily seen to be symmetric. Such a reduction is not possible when the determinant $|S_{ik}|$ vanishes.

It is evident that every solution of the system (53) defines a solution of (55) by means of the transformation $u_i=B_{\alpha i}z_\alpha$. Conversely, if a solution $u_i(x)$ of the equations (55) is known, for a particular value of λ, then the functions

$$z_i(x) = \lambda\int_a^b G_{\beta i}(\xi,x)u_\beta(\xi)d\xi$$

satisfy the relations $u_i=B_{\alpha i}z_\alpha$ and the equations (53). If the boundary value problem (16) equivalent to the system (52) is definitely self-adjoint, every solution of the system (52) defines a solution of the conjugate system (55) by means of the symmetric transformation $u_i=S_{i\beta}y_\beta=B_{\alpha i}\,T_{\alpha\beta}\,y_\beta$, as one may infer from the transformation $z_i=T_{i\beta}\,y_\beta$ relating the solutions of the system (52) and its adjoint (53), or directly by means of the relations (54) and (24).

It is clear then that the theory of a definitely self-adjoint boundary value problem (16) may be regarded as a special case of the theory of a system of linear integral equations (52) whose solutions go over into solutions of the conjugate system (55) by means of a symmetric transformation $u_i=S_{i\beta}\,y_\beta$. From the preceding paragraphs it is evident that such systems have many properties analogous to those of systems whose kernel matrix is symmetric. The symmetric case is the one which arises when the matrix of functions S_{ik} is the identity matrix. It would be interesting to investigate in detail the theory of such systems of linear integral equations.

6. The convergence proof. For the purpose of proving the uniform convergence of the series (36) of Section 3 a number of lemmas are required

* See, for example, Bôcher, *Higher Algebra*, p. 299.

which are analogous to those used for similar purposes in the theory of a single linear integral equation,* and which have for the most part been frequently applied in more generalized form than here given by E. H. Moore in his "general analysis."

It is understood that the functions $S_{ik}(x)$ are continuous and the quadratic form $S_{\alpha\beta} f_\alpha f_\beta \geqq 0$ at every x on the interval ab and for every set of arguments f_i. The functions $f_i(x)$, $g_\alpha(x)$, $y_{i\mu}(x)$ in the following theorems are at least bounded and integrable on the interval ab, and the sets $y_{i\mu}(x)$ ($\mu = 1, 2, \cdots$) are normed and orthogonal in the sense described in Theorem 10. The constants γ_μ are the numbers

$$\gamma_\mu = \int_a^b f_\alpha(x) S_{\alpha\beta}(x) y_{\beta\mu}(x) dx.$$

The first lemma below follows at once from the readily proved equation

$$\int_a^b [f_\alpha - \sum \gamma_\mu y_{\alpha\mu}] S_{\alpha\beta} [f_\beta - \sum \gamma_\nu y_{\beta\nu}] dx = \int_a^b f_\alpha S_{\alpha\beta} f_\beta dx - \sum \gamma_\mu^2, \tag{56}$$

where the sums without range indicated are taken with respect to μ, ν over the same arbitrarily selected set of a finite number of positive integers. The second and third lemmas are immediate consequences of the first one. The integral in the second member of equation (56) is called the norm of the set $f_i(x)$.

LEMMA 1. $\sum \gamma_\mu^2 < \int_a^b f_\alpha S_{\alpha\beta} f_\beta \, dx$.

LEMMA 2. *$\sum_{\mu=1}^{\infty} \gamma_\mu^2$ converges.*

LEMMA 3. *If the norm of the set $g_i(x)$ is zero then*

$$\int_a^b f_\alpha S_{\alpha\beta} g_\beta dx \leqq \int_a^b f_\alpha S_{\alpha\beta} f_\beta dx \int_a^b g_\alpha S_{\alpha\beta} g_\beta dx.$$

Suppose now that the functions $h_i(x, \xi)$ are bounded for all values of x and ξ on the interval ab, and integrable in x on that interval for every fixed ξ. Let δ_μ and $\epsilon_\mu(\xi)$ denote the integrals

$$\delta_\mu = \int_a^b g_\alpha S_{\alpha\beta} y_{\beta\mu} dx, \qquad \epsilon_\mu(\xi) = \int_a^b h_\alpha(x, \xi) S_{\alpha\beta}(x) y_{\beta\mu}(x) dx.$$

---

* See E. Schmidt, pp. 1–4.

LEMMA 4. *The series* $\sum_{\mu=1}^{\infty} \delta_\mu \epsilon_\mu(\xi)$ *converges uniformly on the interval* $a \leqq \xi \leqq b$.

The proof is analogous to one given by Schmidt.* In the sum

$$\sum_{\mu=p}^{p+q} | \delta_\mu \epsilon_\mu(\xi) | = \sum \delta_\rho \epsilon_\rho(\xi) - \sum \delta_\sigma \epsilon_\sigma(\xi)$$

let the index ρ range over those numbers for which $\delta_\rho \epsilon_\rho(\xi)$ is positive for a fixed ξ, and σ over those for which these terms are negative. Then by Lemma 3

$$\begin{aligned} \sum \delta_\rho \epsilon_\rho(\xi) &= \int_a^b h_\alpha(x,\xi) S_{\alpha\beta}(x) \sum \delta_\rho y_{\beta\rho}(x) dx \\ &\leqq \left[\int_a^b h_\alpha S_{\alpha\beta} h_\beta dx \cdot \sum \delta_\mu^2 \right]^{1/2} \\ &\leqq A^{1/2} \left(\sum_{\mu=p}^{\infty} \delta_\mu^2 \right)^{1/2}, \end{aligned}$$

where A is the maximum of the norm of the functions $h_\alpha(x, \xi)$ on the interval $a \leqq \xi \leqq b$. A similar inequality holds for the negative terms, and it follows that

$$\sum_{\mu=p}^{p+n} | \delta_\mu \epsilon_\mu(\xi) | \leqq 2A^{1/2} \left(\sum_{\mu=p}^{\infty} \delta_\mu^2 \right)^{1/2}.$$

Since the series of Lemma 2 converges it is evident that the similar series of terms δ_μ^2 has the same property, and from the last inequality it follows that the series of Lemma 4 converges uniformly on the interval ab.

THEOREM 17. *If the functions* $f_i(x)$ *are solutions of a system of the form*

$$f_i' = A_{i\alpha} f_\alpha + B_{i\alpha} g_\alpha, \qquad s_i(f) = 0$$

then the series $\sum_{\mu=1}^{\infty} \gamma_\mu y_{i\mu}(x)$ *converges uniformly on the interval* ab.

From equations (38) and (52) it follows that the series of the theorem are also expressible in the form

$$\begin{aligned} \sum_{\mu=1}^{\infty} \gamma_\mu y_{i\mu}(x) &= \sum_{\mu=1}^{\infty} \frac{1}{\lambda_\mu} y_{i\mu}(x) \int_a^b g_\alpha S_{\alpha\beta} y_{\beta\mu} dx \\ &= \sum_{\mu=1}^{\infty} \delta_\mu \int_a^b G_{i\alpha}(x,\xi) T_{\alpha\beta}^{-1}(\xi) S_{\beta\gamma}(\xi) y_{\gamma\mu}(\xi) d\xi. \end{aligned}$$

* Pp. 2–4.

For each fixed value of i this is a series of the form whose uniform convergence is stated in Lemma 4.

A LIST OF REFERENCES

(1) E. Schmidt, *Entwickelung willkürlicher Funktionen nach Systemen vorgeschriebener.* Dissertation, Göttingen, 1905.
(2) Bounitzky, Journal de Mathématiques, ser. 6, vol. 5 (1909), p. 65.
(3) Hildebrandt, these Transactions, vol. 19 (1918), p. 94.
(4) E. H. Moore, Bulletin of the American Mathematical Society, vol. 18 (1912), p. 334.
(5) Schur, Mathematische Annalen, vol. 72 (1921), p. 213.
(6) Hurwitz, these Transactions, vol. 22 (1921), p. 526.
(7) Camp, American Journal of Mathematics, vol. 44 (1922), p. 25.
(8) Carmichael, American Journal of Mathematics, vol. 43 (1921), pp. 69, 232; vol. 44 (1922), p. 129; Bulletin of the American Mathematical Society, vol. 28 (1922), p. 200.
(9) Birkhoff-Langer, Proceedings of the American Academy of Arts and Sciences, vol. 58 (1923), p. 51. See also their abstract, Bulletin of the American Mathematical Society, vol. 27 (1922), p. 236.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ON THE THEORY OF INTEGRAL EQUATIONS WITH DISCONTINUOUS KERNELS*

BY

RUDOLPH E. LANGER

Chapter 1. Introduction

This paper is concerned with the theory of the integral equation whose salient characteristic is that its kernel $K(x, \xi)$ is discontinuous along the line $\xi = x$. It is closely related with the theory of the integral equation characterized by discontinuities in the partial derivatives of the kernel for $\xi = x$, of which more specific mention is made below.

The point of departure for the consideration involved is in each case furnished by the familiar fact of the equivalence of a differential system composed of an equation

$$p_0 \frac{d^n u}{dx^n} + \cdots + p_n u = \lambda u,$$

and homogeneous boundary conditions, with the integral equation

$$u(x) = \lambda \int_a^b G(x, \xi) u(\xi) d\xi,$$

in which the kernel is the Green's function of the reduced differential system. The existing theory of differential boundary problems guarantees the existence of characteristic values of λ and corresponding functions $u(x)$ for the differential system in a broad class of cases. The same is true, therefore, for the equivalent integral equation, and viewed from the standpoint of the theory of integral equations these facts must be attributable to the peculiarities of the Green's function which serves as a kernel.

Fixing the attention for the moment on the case in which the differential system is of the second order and self-adjoint, the properties of the kernel which are of particular interest in this connection are (1), its symmetry in its arguments, and (2), the finite non-vanishing discontinuities in its first partial derivatives for $\xi = x$. The extent to which symmetry of the kernel serves as a basis for a theory of integral equations is shown by the theory of Hilbert and Schmidt. The question of the extent to which the second

* Presented to the Society, January 1, 1926; received by the editors January 21 and March 26, 1926.

property mentioned serves as such a basis has remained open. It appears to have been touched upon first by Birkhoff in 1906 in an unpublished discussion of the equation with kernel both symmetric and discontinuous.*

The development of a theory based only on the discontinuity of the first partial derivatives of the kernel was taken up by Mrs. Eleanor P. Brown† in a thesis prepared under the direction of Birkhoff and presented at Radcliffe College in 1921. This theory will be presented shortly in a joint paper by Mrs. Brown and the author.

If the differential system under consideration is of the first order the discontinuity along the line $\xi = x$ occurs in the Green's function itself. It is this property which is taken as the distinguishing feature of the kernel of the integral equation to which the present paper is devoted. While the analytic details which are of interest differ in many respects, the methods used here are in the main parallel to those of the paper referred to above which is to embody Mrs. Brown's thesis. They are similar also in many respects to the methods employed by Birkhoff and the author in a treatment of differential boundary problems.‡

The content of the paper may be roughly summarized as follows. In Chapter 2 the hypotheses on the kernel of the given equation are enunciated and a change of variables is deduced which serves to normalize the discontinuities in the kernel and its first partial derivatives. In Chapter 3 the integral equation is transformed into a system composed of an integro-differential equation and a boundary condition, and in Chapter 4 the equivalence of this system with the given equation is established. Chapters 5, 6, and 7 are concerned with the integro-differential equation alone (without the boundary condition), and it is shown that the equation is possessed of a solution for all values of the parameter which are sufficiently large. The asymptotic form of this solution in any right-hand or left-hand half-plane is there deduced.

In Chapter 8 the boundary condition is introduced and by means of it the existence, under certain restrictions, of infinitely many characteristic values for the parameter is proved and the asymptotic form of these values is obtained. In Chapters 9, 10, and 11, the normalized asymptotic forms of the characteristic functions for the given integral equation and its associated

---

* Bulletin of the American Mathematical Society, vol. 13 (1906), p. 62.

† Formerly Miss Eleanor Pairman.

‡ G. D. Birkhoff and R. E. Langer, *Boundary problems and developments associated with a system of ordinary linear differential equations of the first order*, Proceedings of the American Academy of Arts and Sciences, vol. 58, pp. 51–128; April, 1923

equation are developed, and it is proved that when the given equation has only simple characteristic values the set of characteristic functions is closed. Lastly in Chapter 12 the expansion of an arbitrary function in a series of the solutions is considered. It is shown that for the expansion of any function $f(x)$ which is integrable in the sense of Lebesgue the solutions of the integral equation have essentially the same properties as the solutions of a related system composed of a differential equation of the first order and a homogeneous boundary condition. At the end points of the interval, and only there, do the two expansions behave in dissimilar fashion. The related differential system in question here is a special case of a system previously studied by the author.*

In the appendix a simple example of an integral equation to which the developed theory applies is given, and by computed results various features of the theory are illustrated.

Chapter 2. The normalization of the equation

1. **The given equation.** We consider given the equation

$$y(t) = \rho\int_\alpha^\beta \Gamma(t,\tau)y(\tau)d\tau, \tag{1}$$

in which ρ is a complex parameter. The kernel is real and satisfies the following conditions:
(i) that it is possessed of partial derivatives to those of order $n \geqq 1$ inclusive, these partial derivatives being continuous in the open regions

$$R'\begin{cases}\alpha \leqq \tau < t\\ \alpha \leqq t \leqq \beta\end{cases} \quad\text{and}\quad R''\begin{cases}t < \tau \leqq \beta\\ \alpha \leqq t \leqq \beta,\end{cases}$$

and approaching in each of these regions a finite limiting value at every point of the boundary $\tau = t$;
(ii) that

$$\Gamma(t,\tau)\Big]_{\tau=t-}^{\tau=t+} = \varphi(t) \neq 0.$$

2. **The derivatives of $\varphi(t)$.** The differentiability of the function $\varphi(t)$ defined in (ii) may be derived from the facts which we formulate as follows.

---

* *Developments associated with a boundary problem*, these Transactions, vol. 25 (1923), pp. 155–172.

LEMMA. *If $F(t, \tau)$ is any function possessed of the properties* (i) *above, then*

$$\text{(a)} \qquad \frac{dF(t,t+)}{dt} = F_t(t,t+) + F_\tau(t,t+),$$

$$\text{(b)} \qquad \frac{dF(t,t-)}{dt} = F_t(t,t-) + F_\tau(t,t-).$$

Let $t_2 > t_1$ be any two points of the interval (α, β). Then

$$\frac{F(t_2,t_2+) - F(t_1,t_1+)}{t_2 - t_1} = \frac{1}{\Delta t}\{F(t_2,t_2+) - F(t_1,t_2) + F(t_1,t_2) - F(t_1,t_1+)\},$$

where $\Delta t = t_2 - t_1$. By the law of the mean the right hand member of this equals

$$F_t(t_2 - \theta_1\Delta t, t_2) + F_\tau(t_1, t_2 - \theta_2\Delta t), \qquad 0 < \theta_i < 1.$$

Since this approaches the right hand member of (a) as a limit when either t_1 or t_2 is taken at t and $\Delta t \to 0$ the proof of statement (a) is complete.

The proof of (b) is precisely similar.

Since the kernel $\Gamma(t, \tau)$ and its partial derivatives to those of order $(n-1)$ are of the type of $F(t, \tau)$ of the lemma, we may differentiate the right hand member of the relation

$$\varphi(t) = \Gamma(t,t+) - \Gamma(t,t-)$$

n times. This, together with the continuity of the resulting terms on the right, establishes the existence and continuity of $d^n\varphi(t)/dt^n$ on the interval (α, β).

3. **The change of variables.** By hypothesis the function $\varphi(t)$ maintains its sign on (α, β). With a proper distribution of the constant factors between ρ and Γ it will follow that $\varphi(t) > 0$. We assume such a distribution. Then the continuous functions

$$\text{(2)} \qquad x = \int_\alpha^t \varphi(t)dt, \qquad \xi = \int_\alpha^\tau \varphi(\tau)d\tau$$

are monotonic increasing and may be used as new independent variables. Introducing them in equation (1) and denoting by $\bar{f}(x, \xi)$ the function into which $f(t, \tau)$ is transformed by the inverse of (2), we obtain the form

$$\text{(3)} \qquad \bar{y}(x) = \lambda \int_a^b \Omega(x,\xi)\bar{y}(\xi)d\xi,$$

in which

$$(4)\qquad \lambda = -\rho, \quad a = 0, \quad b = \int_\alpha^\beta \varphi(\tau)d\tau,$$

$$\Omega(x,\xi) = -\frac{\bar{\Gamma}(x,\xi)}{\bar{\varphi}(\xi)}\cdot$$

It is readily verified that in the regions

$$R_1\begin{cases} a \leqq \xi < x \\ a \leqq x \leqq b \end{cases} \quad \text{and} \quad R_2\begin{cases} x < \xi \leqq b \\ a \leqq x \leqq b, \end{cases}$$

and on the boundary $\xi = x$, $\Omega(x, \xi)$ is possessed of properties analogous to those under (i) above. Moreover

$$\Omega(x,\xi)\Big]_{\xi=x-}^{\xi=x+} \equiv -1.$$

Defining $\psi(x)$ now by the relation

$$\Omega_x(x,\xi)\Big]_{\xi=x-}^{\xi=x+} = \psi(x),$$

and introducing in (3) as a new dependent variable the function

$$u(x) = \bar{y}(x)e^{\int_a^x \psi(x)dx},$$

we obtain as the final form of the equation

$$(5)\qquad u(x) = \lambda\int_a^b K(x,\xi)u(\xi)d\xi,$$

where

$$K(x,\xi) = \Omega(x,\xi)e^{\int_\xi^x \psi(x)dx}.$$

This kernel $K(x, \xi)$ and its partial derivatives are undefined on the line $\xi = x$. We shall complete their definitions by designating them to have in the points of this line their limiting values as these points are approached in the region R_1.

4. **The normalized equation.** The equation (5) will be said to represent the normal form because of the following characteristics which its kernel is found to possess.

(A) The function $K(x, \xi)$ is possessed of partial derivatives to those of order $n \geqq 1$ inclusive, which are continuous in the region R_2 and in the closed region

$$R_1'\begin{cases} a \leqq \xi \leqq x \\ a \leqq x \leqq b\,; \end{cases}$$

(B) $$K(x,\xi)\Big]_{\xi=x}^{\xi=x+} \equiv -1;$$

(C) $$K_x(x,\xi)\Big]_{\xi=x}^{\xi=x+} \equiv 0.$$

It may be observed that the differentiation of (B) and the application of (C) yield the further relation

(D) $$K_\xi(x,\xi)\Big]_{\xi=x}^{\xi=x+} \equiv 0.$$

THEOREM 1. *Every integral equation of the type* (1) *with a kernel possessed of the properties* (i) *and* (ii) *above may, by a suitable change of variables, be transformed into an equation of the same type with a kernel possessed of the properties* (A), (B), (C) *and* (D).

CHAPTER 3. TRANSFORMATION OF THE INTEGRAL EQUATION INTO AN INTEGRO-DIFFERENTIAL SYSTEM

5. **The auxiliary differential system.** We consider in connection with the normalized equation the differential system

(6) $$y'(x) = 0, \qquad \mu y(a) + \nu y(b) = 0.$$

The constants μ and ν are parameters, the choice of which is arbitrary, subject to restrictions to be imposed at certain points in the subsequent theory. In each case, however, only discrete values of the ratio μ/ν will be excluded.

To begin, let μ and ν be chosen so that

(7) $$\mu/\nu \neq -1.$$

Then system (6) is incompatible and possesses a Green's function $G(x, \xi)$. This function and its derivatives are undefined on the line $\xi = x$ where $G(x, \xi)$ is discontinuous. We shall complete their definitions in the region R, composed of R_2 and R_1', by designating them to have in the points of the line $\xi = x$ their limiting values as these points are approached in the region R_1. We may then enumerate their characteristics as follows.

I. The functions $G(x, \xi)$, $G_x(x, \xi)$, and $G_\xi(x, \xi)$ are continuous in R_2 and R_1';

II. $$G(x,\xi)\Big]_{\xi=x}^{\xi=x+} \equiv -1;$$

III. $$G_x(x,\xi) \equiv G_\xi(x,\xi) \equiv 0;$$

IV. $$\mu G(a,\xi) + \nu G(b,\xi) \equiv 0;$$

V. $$\nu G(x,a) + \mu G(x,b) \equiv 0;$$

VI. the solution of the system

$$\omega'(x) = f(x),$$
$$\mu\omega(a) + \nu\omega(b) = 0$$

is given by

$$\omega(x) = \int_a^b G(x,t)f(t)dt\,;$$

VII. $$G(x,\xi) = -H(\xi,x),$$

where $H(x, \xi)$ is the Green's function of the system adjoint to (6).

6. **The construction of $\boldsymbol{K}(x, \xi)$.** It will be observed that $G(x, \xi)$ and the kernel $K(x, \xi)$ of the normalized equation (5), and their respective first partial derivatives, have the same discontinuities in the region R. We shall construct with the use of $K(x, \xi)$ another function $\boldsymbol{K}(x, \xi)$ which maintains these characteristics, and in addition shares with $G(x, \xi)$ its properties IV and V.

We assume now that for the given equation

(iii) $$n \geqq 2,$$

and that in the normalized form

(iv) $$|K(a,b)| + |K(a,a+) + K(b,b)| + |K(b,a)| \neq 0.$$

Then μ and ν may be so chosen, subject to previous restriction, that

(8) $$\Delta = \mu^2K(a,b) + \mu\nu[K(a,a+) + K(b,b)] + \nu^2K(b,a) \neq 0.$$

We assume such a choice, and set

(9) $$\begin{aligned}\nu K(x,a) + \mu K(x,b) &= W(x),\\ \mu K(a,\xi) + \nu K(b,\xi) &= V(\xi).\end{aligned}$$

Clearly these functions $W(x)$ and $V(\xi)$ are possessed of continuous derivatives to those of order n on the interval (a, b). The function $\boldsymbol{K}(x, \xi)$, defined by the formula

(10) $$\boldsymbol{K}(x,\xi) = \frac{1}{\Delta}\begin{vmatrix} K(x,\xi) & W(x) \\ V(\xi) & \Delta \end{vmatrix},$$

is, therefore, found to possess the properties (A) to (D) of the kernel $K(x, \xi)$. From the relations

(11) $$\begin{aligned}\mu W(a) + \nu W(b) &= \Delta,\\ \nu V(a) + \mu V(b) &= \Delta,\end{aligned}$$

it follows further that

$$\mu K(a,\xi) + \nu K(b,\xi) \equiv 0, \tag{12}$$
$$\nu K(x,a) + \mu K(x,b) \equiv 0.$$

7. **The relation between** $K(x, \xi)$ **and** $G(x, \xi)$. If x is regarded for the moment as a parameter, the function

$$\omega(x,\xi) = K(x,\xi) - G(x,\xi)$$

is continuous together with its first derivative in ξ. Since by III, V, and (12) it is also a solution of the differential system

$$\frac{\partial\omega(x,\xi)}{\partial\xi} = K_\xi(x,\xi),$$

$$\nu\omega(x,a) + \mu\omega(x,b) = 0,$$

it is uniquely determined as such, and is by VI expressible in terms of the Green's function $H(x, \xi)$. Introducing $G(x, \xi)$ by virtue of VII, we may write therefore

$$K(x,\xi) - G(x,\xi) = -\int_a^b K_t(x,t)G(t,\xi)dt. \tag{13}$$

This relation, if ξ is now looked upon as the parameter, is in form an integral equation for $G(x, \xi)$ as a function of x.

8. **The solvability of the relation between** $K(x,\xi)$ **and** $G(x, \xi)$. A sufficient condition for the solvability of equation (13) for $G(x, \xi)$ is that the Fredholm determinant D for the kernel $K_\xi(x, \xi)$ shall differ from zero. This kernel, and hence also D, depends upon the parameters μ and ν. We shall show that under the assumption, which we now make, that the given equation is such that

$$D \not\equiv 0 \text{ in } \mu \text{ and } \nu, \tag{v}$$

it is always possible to choose these parameters subject to previous restrictions, so that $D\neq 0$.

We set

$$K_\xi(x,\xi) = E(x,\xi),$$
$$K_\xi(x,\xi) = \mathcal{E}(x,\xi)$$

and denote respectively by D and $D(x, \xi)$ the Fredholm determinant and first minor of the kernel $E(x, \xi)$, and by $\mathcal{D}$ and $\mathcal{D}(x, \xi)$ the corresponding expressions for the kernel $\mathcal{E}(x, \xi)$. We have*

$$D = 1 + \sum_{n=1}^{\infty} d_n\ , \quad D(\xi_1,\xi_2) = E(\xi_1,\xi_2) + \sum_{n=1}^{\infty} d_n(\xi_1,\xi_2),$$

* Bôcher, *An Introduction to Integral Equations*, Cambridge University Press, 1909, p. 32.

where, if we omit writing the arguments and indicate them only by their subscripts, thus: $E(\xi_i,\xi_j)=E_{ij}$, then

$$d_n = \frac{(-1)^n}{n!}\int_a^b \cdots \int_a^b \begin{vmatrix} E_{11} & \cdots & E_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ E_{n1} & \cdots & E_{nn} \end{vmatrix} d\xi_1 \cdots d\xi_n,$$

$$d_n(\xi_1,\xi_2) = \frac{(-1)^n}{n!}\int_a^b \cdots \int_a^b \begin{vmatrix} E_{12} & E_{13} & \cdots & E_{1,n+2} \\ E_{32} & E_{33} & \cdots & E_{3,n+2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ E_{n+2,2} & & \cdots & E_{n+2,n+2} \end{vmatrix} d\xi_3 \cdots d\xi_{n+2}.$$

We shall omit writing down the analogous formulas for $\mathcal{D}$ and $\mathcal{D}(\xi_1, \xi_2)$, since they may be obtained by replacing in those above d_n, $d_n(\xi_1, \xi_2)$ and E respectively by $\mathfrak{d}_n$, $\mathfrak{d}_n(\xi_1, \xi_2)$ and $\mathcal{E}$.

Consider the term d_n. By (10) we have

$$E(x,\xi) = \mathcal{E}(x,\xi) + \varphi(x)\psi(\xi),$$

where

$$\varphi(x)\psi(\xi) = \frac{W(x)V(\xi)}{\Delta}\cdot$$

Substituting this form and expanding the resulting determinant into a sum of determinants with monomial elements we obtain

$$d_n = \frac{(-1)^n}{n!}\int_a^b \cdots \int_a^b \begin{vmatrix} \mathcal{E}_{11} & \cdots & \mathcal{E}_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \mathcal{E}_{n1} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1 \cdots d\xi_n$$

$$+ \frac{(-1)^n}{n!}\int_a^b \cdots \int_a^b \sum_{i=1}^{n} \begin{vmatrix} \mathcal{E}_{11} & \cdots & \mathcal{E}_{1,i-1} & \varphi_1\psi_i & \mathcal{E}_{1,i+1} & \cdots & \mathcal{E}_{1n} \\ : & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ : & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ \mathcal{E}_{n1} & \cdots & \mathcal{E}_{n,i-1} & \varphi_n\psi_i & \mathcal{E}_{n,i+1} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1 \cdots d\xi_n.$$

Of the terms on the right the first is simply d_n. The remaining sum we shall reduce by the following manipulation. In the determinant of the ith term of the sum, let the ith row and column be shifted into first places, and let the arguments be renamed as follows:

ξ_k to become ξ_{k+1}, for $k=1, 2, \cdots, (i-1)$;

ξ_i to become ξ_1;

the others to remain unchanged. Since these arguments are all variables of integration this change in their designation amounts merely to a change in the order of integration, and this is immaterial. The terms of the sum have thus been made identical, and hence the entire expression reduces to

$$\frac{(-1)^n}{n!}\int_a^b\cdots\int_a^b n\psi_1 \begin{vmatrix} \varphi_1 & \mathcal{E}_{12} & \cdots & \mathcal{E}_{1n} \\ \varphi_2 & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \varphi_n & \mathcal{E}_{n2} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1\cdots d\xi_n.$$

This we expand by the elements of the first column to obtain

$$\frac{(-1)^n}{(n-1)!}\int_a^b\cdots\int_a^b \psi_1\varphi_1 \begin{vmatrix} \mathcal{E}_{22} & \cdots & \mathcal{E}_{2n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \mathcal{E}_{n2} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1\cdots d\xi_n$$

$$+\frac{(-1)^n}{(n-1)!}\int_a^b\cdots\int_a^b \psi_1 \sum_{j=2}^{n}(-1)^{j-1}\varphi_j \begin{vmatrix} \mathcal{E}_{12} & \cdots & \mathcal{E}_{1n} \\ \cdot & \cdots & \cdot \\ \mathcal{E}_{j-1,2} & \cdots & \cdot \\ \mathcal{E}_{j+1,2} & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \mathcal{E}_{n2} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1\cdots d\xi_n.$$

Observing that the first term of this is

$$-\int_a^b \varphi(\xi_1)\psi(\xi_1)d_{n-1}d\xi_1,$$

we proceed to a further reduction of the remaining sum. Let the $(j-1)$th column in the respective term (i.e., that in which j is the second subscript) be shifted into first place and let the arguments again be renamed so that

ξ_k becomes ξ_{k+1}, for $k=1, 2, \cdots, (j-1)$,

ξ_j becomes ξ_1,

and the others remain unchanged. This again makes the terms of the sum identical, and the expression takes the form

$$\frac{(-1)^{n-1}}{(n-1)!}\int_a^b \cdots \int_a^b (n-1)\varphi_1\psi_2 \begin{vmatrix} \mathcal{E}_{21} & \mathcal{E}_{23} & \cdots & \mathcal{E}_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \mathcal{E}_{n1} & \mathcal{E}_{n3} & \cdots & \mathcal{E}_{nn} \end{vmatrix} d\xi_1 \cdots d\xi_n.$$

This is precisely

$$-\int_a^b\int_a^b \varphi(\xi_1)\psi(\xi_2)d_{n-2}(\xi_2,\xi_1)d\xi_1 d\xi_2.$$

Collecting our results, then, and summing the d_n to obtain D, we find that

$$D = \mathcal{D} - \frac{1}{\Delta}\int_a^b W(\xi_1)\left\{V'(\xi_1)\mathcal{D} + \int_a^b V'(\xi_2)\,\mathcal{D}(\xi_2,\xi_1)d\xi_2\right\}d\xi_1.$$

By use of Fredholm's relation

$$\mathcal{E}(x,\xi_1)\mathcal{D} + \int_a^b \mathcal{E}(x,\xi_2)\,\mathcal{D}(\xi_2,\xi_1)d\xi_2 = \mathcal{D}(x,\xi_1),$$

and the formula

$$V'(\xi_1) = \mu\mathcal{E}(a,\xi_1) + \nu\mathcal{E}(b,\xi_1),$$

which follows from (9), the final form

$$D = \mathcal{D} - \frac{1}{\Delta}\int_a^b W(\xi_1)\left\{\mu\mathcal{D}(a,\xi_1) + \nu\mathcal{D}(b,\xi_1)\right\}d\xi_1$$

is obtained. We observe now that in this relation $\mathcal{D}$ and $\mathcal{D}(x, \xi)$ are free from the parameters μ and ν, since they depend only on $\mathcal{E}(x, \xi)$. It follows that D is rational in μ and ν. The condition $D \not\equiv 0$ in μ, ν is found directly to take the explicit form

(va) $$|\theta(a,b)| + |\theta(a,a+) + \theta(b,b)| + |\theta(b,a)| \neq 0,$$

where

$$\theta(x,y) = K(x,y)\mathcal{D} - \int_a^b \mathcal{D}(x,\xi)K(\xi,y)d\xi,$$

and $\mathcal{D}$ and $\mathcal{D}(x,\xi)$ are the Fredholm determinant and first minor of the kernel $K_\xi(x, \xi)$. Except in the case that this inequality fails, the parameters μ and ν can be chosen so that $D \neq 0$.

The method of manipulation used on D can be employed equally well in the analysis of $D(\xi_1, \xi_2)$. It may be shown in this way that $D(\xi_1, \xi_2)$ is also rational in μ and ν.

With μ and ν chosen so that $D\neq 0$, the existence of the kernel $F(x, \xi)$ reciprocal to $E(x, \xi)$ is assured. Writing the relation (13) then in "solved" form we have

$$G(x,\xi) = K(x,\xi) - \int_a^b F(x,t)K(t,\xi)dt. \tag{14}$$

9. **Characteristics of** $F(x, \xi)$. Since the kernel $E(x, \xi)$ is continuous in R, it follows that $F(x, \xi)$ is likewise continuous in R. The behavior of the partial derivatives of $F(x, \xi)$ may be determined as follows. Differentiating Fredholm's identity

$$E(x,\xi) + F(x,\xi) \equiv \int_a^b E(x,t)F(t,\xi)dt \equiv \int_a^b F(x,t)E(t,\xi)dt, \tag{15}$$

we obtain

$$E_x(x,\xi) + F_x(x,\xi) \equiv \int_a^b E_x(x,t)F(t,\xi)dt,$$

$$E_\xi(x,\xi) + F_\xi(x,\xi) \equiv \int_a^b F(x,t)E_\xi(t,\xi)dt.$$

Since the right hand members of these relations are obviously continuous it follows that $F_x(x, \xi)$ and $F_\xi(x, \xi)$ are continuous in R_2 and R_1', and hence that

$$\begin{aligned} F_x(x,\xi)\Big]_{\xi=x}^{\xi=x+} &\equiv - E_x(x,\xi)\Big]_{\xi=x}^{\xi=x+}, \\ F_\xi(x,\xi)\Big]_{\xi=x}^{\xi=x+} &\equiv - E_\xi(x,\xi)\Big]_{\xi=x}^{\xi=x+}. \end{aligned} \tag{16}$$

In similar manner the higher partial derivatives may be considered. It will be found that $F(x, \xi)$, like $E(x, \xi)$, is possessed in R_2 and R_1' of continuous partial derivatives to those of order $(n-1)$.

From the formula

$$F(x,\xi) = -\frac{D(x,\xi)}{D},$$

it is seen further that $F(x, \xi)$ is rational in the parameters μ and ν. From Fredholm's identity above and the relation

$$\mu E(a,\xi) + \nu E(b,\xi) \equiv 0, \tag{17a}$$

obtained by differentiation from (12), we draw lastly that

$$\mu F(a,\xi) + \nu F(b,\xi) \equiv 0. \tag{17b}$$

10. The transformation of the integral equation. Let $u(x)$ represent, now, any solution of the normalized equation (5). Multiplying (14) by $\lambda u(\xi)$ and integrating with respect to ξ, we obtain

$$\lambda\int_a^b G(x,\xi)u(\xi)d\xi = \lambda\int_a^b K(x,\xi)u(\xi)d\xi - \int_a^b F(x,t)\lambda\int_a^b K(t,\xi)u(\xi)d\xi \cdot dt. \tag{18}$$

In this the integral which occurs in both terms on the right may be evaluated by multiplying (10) by $\lambda u(\xi)$, integrating, and utilizing (5). With the result so obtained and with the abbreviation

$$L(u) = \mu u(a) + \nu u(b), \tag{19}$$

the equation (18) reduces to the form

$$\begin{aligned}\lambda\int_a^b G(x,\xi)u(\xi)d\xi = u(x) - L(u)\,\frac{W(x)}{\Delta}\\ - \int_a^b F(x,t)\left\{u(t) - L(u)\frac{W(t)}{\Delta}\right\}dt.\end{aligned} \tag{20}$$

We shall abbreviate this result by setting

$$\Phi(x) = \frac{1}{\Delta}\left[W(x) - \int_a^b F(x,t)W(t)dt\right]. \tag{21}$$

The function $\Phi(x)$ so defined is continuous and has a continuous $(n-1)$th derivative on (a, b). For subsequent use we observe that

$$\mu\Phi(a) + \nu\Phi(b) = 1. \tag{22}$$

Introducing $\Phi(x)$ into (20) we may write that relation

$$u(x) - \lambda\int_a^b G(x,\xi)u(\xi)d\xi = L(u)\Phi(x) + \int_a^b F(x,t)u(t)dt. \tag{23}$$

By differentiation this yields finally the equation

$$u'(x) - \lambda u(x) = L(u)\varphi(x) + \int_a^b f(x,t)u(t)dt,$$

where we have set

$$\begin{aligned}\varphi(x) &= \Phi'(x),\\ f(x,\xi) &= F_x(x,\xi).\end{aligned} \tag{24}$$

Lastly adjoining to the equation thus derived a second equation—obviously satisfied by every solution of (5)—we may formulate the result as follows.

THEOREM 2. *If the kernel of the normalized integral equation* (5) *satisfies the conditions* (iii),(iv),(v), *above, then every solution of the equation is also a solution of the related integro-differential system*

$$(25)\quad \begin{aligned} &\text{(a)} \qquad u'(x) - \lambda u(x) = L(u)\varphi(x) + \int_a^b f(x,t)u(t)dt, \\ &\text{(b)} \qquad L(u) = \lambda \int_a^b V(t)u(t)dt. \end{aligned}$$

The functions $V(t)$ and $L(u)$ involved here are given respectively by (9) and (19) above. We remark that the function $\varphi(x)$ is continuous together with its derivatives to that of order $(n-2)$ on (a, b), while $f(x, \xi)$ is continuous with its partial derivatives to those of order $(n-2)$ in R_2 and R_1'.

11. The associated integral equation. Deductions analogous to those above may also be made for the integral equation

$$v(x) = \lambda \int_a^b K(\xi, x)v(\xi)d\xi,$$

associated with equation (5). To accomplish this most easily we shall set $-K(\xi, x) = \bar{K}(x, \xi)$ and write the equation in the form

$$(\bar{5}) \qquad v(x) = -\lambda \int_a^b \bar{K}(x,\xi)v(\xi)d\xi,$$

which is readily found to be its normal form. The deductions already made become applicable to equation ($\bar{5}$), then, if λ is replaced by $-\lambda$ and if the functions involved are taken to spring from the kernel $\bar{K}(x, \xi)$ rather than from $K(x, \xi)$. When this is the case we shall indicate it by superscribing the various functional symbols with a bar.

The relation between the two developments is more easily followed if in passing from the treatment of the given equation to that of the associated equation the parameters μ and ν are interchanged. We shall suppose this done. Since $\bar{\Delta} \equiv -\Delta$, then, the condition $\bar{\Delta} \not\equiv 0$ has already been met.

It is, however, necessary to impose upon the kernel of the given equation the condition that

$$(\bar{\text{v}}) \qquad \bar{D} \not\equiv 0 \text{ in } \nu \text{ and } \mu.$$

The explicit form of this condition as it is obtained from (va) is that

$$(\overline{\text{va}}) \qquad |\bar{\theta}(a,b)| + |\bar{\theta}(a+,a) + \bar{\theta}(b,b)| + |\bar{\theta}(b,a)| \neq 0,$$

where

$$\bar{\theta}(x,y) = K(x,y)\overline{\mathfrak{D}} - \int_a^b \overline{\mathfrak{D}}(y,\xi)K(x,\xi)d\xi,$$

and $\overline{\mathfrak{D}}$ and $\overline{\mathfrak{D}}(x,\xi)$ are the Fredholm determinant and first minor of the kernel $-K_\xi(\xi,x)$.

When condition $(\bar{\text{v}})$ is satisfied, every solution of the equation $(\bar{5})$ is found to be a solution of the system

$$(\overline{25}) \quad \text{(a)} \qquad v'(x) + \lambda v(x) = \bar{L}(v)\bar{\varphi}(x) + \int_a^b \bar{f}(x,t)v(t)dt,$$

$$\text{(b)} \qquad \bar{L}(v) = -\lambda\int_a^b \bar{V}(t)v(t)dt.$$

In this

$$\bar{L}(v) = \nu v(a) + \mu v(b),$$

and to obtain the values of $\bar{f}(x,\xi)$ and $\bar{\varphi}(x)$ in terms of familiar symbols we may proceed as follows.

By direct substitution it is found that

$$\overline{W}(x) \equiv -V(x), \qquad \bar{V}(\xi) \equiv -W(\xi),$$

$$\overline{K}(x,\xi) \equiv -K(\xi,x), \qquad \bar{E}(x,\xi) \equiv -K_\xi(\xi,x).$$

Substituting this value of $\bar{E}(x,\xi)$ in Fredholm's identity

$$\bar{E}(x,\xi) + \bar{F}(x,\xi) = \int_a^b \bar{E}(x,t)\bar{F}(t,\xi)dt$$

and differentiating with respect to x we obtain

$$-K_{\xi x}(\xi,x) + \bar{f}(x,\xi) \equiv -\int_a^b K_{tx}(t,x)\bar{F}(t,\xi)dt.$$

In this we may replace $K_{\xi x}(\xi,x)$ by its equivalent $E_\xi(\xi,x)$. Then further integrating the right hand member of the equation by parts we find that

$$-E_\xi(\xi,x) + \bar{f}(x,\xi) = -E(b,x)\bar{F}(b,\xi) + E(a,x)\bar{F}(a,\xi) + \int_a^b E(t,x)\bar{f}(t,\xi)dt.$$

Because of (17a) and the formula analogous to (17b), however, the sum of the first two terms on the right vanishes. Hence we have the relation

$$E_\xi(\xi,x) - \bar{f}(x,\xi) = -\int_a^b \bar{f}(t,\xi)E(t,x)dt.$$

On the other hand the interchange of the arguments in (15) and the subsequent differentiation of that relation yields

$$E_\xi(\xi,x) + f(\xi,x) = \int_a^b f(\xi,t)E(t,x)dt.$$

A comparison of the two results shows that

$$\bar{f}(x,\xi) = -f(\xi,x).$$

With this result and the formulas already noted above we find readily

$$\bar{\varphi}(x) = \frac{1}{\Delta}\left[V'(x) + \int_a^b V(t)f(t,x)dt\right]. \tag{26a}$$

Lastly we observe that

$$\bar{f}(x,\xi)\Big]_{\xi=x-}^{\xi=x+} = f(x,\xi)\Big]_{\xi=x-}^{\xi=x+}. \tag{26b}$$

Chapter 4. The equivalence of the integro-differential system and the integral equation

12. The transformation of equation (25a). To establish the equivalence of system (25) and equation (5) it remains to show the converse of the theorem of the preceding section, namely that every solution of system (25) is also a solution of equation (5). To do this we shall retrace the steps taken above and so deduce the equation (5) from the system.

Consider first the differential system

$$y'(x) = \varphi(x),$$
$$\mu y(a) + \nu y(b) = 1.$$

Because of (7) it is possessed of a unique continuous solution. By (22) and (24), however, $\Phi(x)$ is such a solution. Hence the system serves to determine $\Phi(x)$ uniquely in terms of $\varphi(x)$. In similar fashion because of (17b) and (24) the system

$$\begin{aligned} &\frac{\partial\omega(x,\xi)}{\partial x} = f(x,\xi), \\ &\mu\omega(a,\xi) + \nu\omega(b,\xi) = 0 \end{aligned} \tag{27}$$

serves to determine $F(x, \xi)$ uniquely in terms of $f(x, \xi)$.

Let $u(x)$ now be any solution of equation (25a), and with the functions $\Phi(x)$ and $F(x, \xi)$ determined construct

$$\sigma(x) = u(x) - \lambda \int_a^b G(x,\xi)u(\xi)d\xi - \int_a^b F(x,\xi)u(\xi)d\xi - L(u)\Phi(x).$$

This function is obviously continuous. Moreover, it is a solution of system (6), for we find on the one hand by differentiation the relation

$$\sigma'(x) = u'(x) - \lambda u(x) - \int_a^b f(x,\xi)u(\xi)d\xi - L(u)\varphi(x),$$

of which the right hand member vanishes by (25a), while it follows on the other hand from property IV of $G(x, \xi)$ and from (17b) and (22) that

$$\mu\sigma(a) + \nu\sigma(b) = 0.$$

The system (6) is incompatible, however. Hence $\sigma(x) \equiv 0$, that is, $u(x)$ satisfies equation (23).

The function $F(x, \xi)$ was originally derived in Chapter 3 as a reciprocal kernel. Hence it is itself possessed of a reciprocal, which is, moreover, precisely $E(x, \xi)$ of the preceding chapter. The integral equations

$$(28) \qquad \begin{aligned} \theta(x) &= \Phi(x) + \int_a^b F(x,t)\theta(t)dt, \\ \omega(x,\xi) &= G(x,\xi) + \int_a^b F(x,t)\omega(t,\xi)dt \end{aligned}$$

are, therefore, uniquely solvable, and since by (21) and (14) we have as solutions respectively $W(x)/\Delta$ and $K(x, \xi)$, these functions are uniquely determinable. We have at our disposal, therefore, the relations (14) and (21) satisfied by them, while the substitution of (21) in (23) yields further the relation (20).

Let (14) be written, now, in "solved" form, thus:

$$K(x,\xi) = G(x,\xi) - \int_a^b E(x,t)\, G(t,\xi)dt.$$

From this we obtain upon multiplying it by $\lambda u(\xi)$ and integrating it with respect to ξ

$$\lambda \int_a^b K(x,\xi)u(\xi)d\xi = \lambda \int_a^b G(x,\xi)u(\xi)d\xi - \int_a^b E(x,t) \cdot \lambda \int_a^b G(t,\xi)u(\xi)d\xi dt.$$

We shall eliminate from this the quantity

$$\lambda\int_a^b G(x,\xi)u(\xi)d\xi,$$

by substituting for it its value as given by (20). In this way we find, upon collecting terms, that

$$\lambda\int_a^b K(x,\xi)u(\xi)d\xi = u(x) - L(u)\frac{W(x)}{\Delta}$$
$$-\int_a^b\left\{F(x,\xi) + E(x,\xi) - \int_a^b E(x,t)F(t,\xi)dt\right\}\left\{u(\xi) - L(u)\frac{W(\xi)}{\Delta}\right\}d\xi.$$

This reduces, however, for by Fredholm's identity (15) the first factor in the integrand on the right vanishes. It follows that every solution of (25a) satisfies also the equation

$$u(x) = \lambda\int_a^b K(x,\xi)u(\xi)d\xi + L(u)\frac{W(x)}{\Delta}\cdot \tag{29}$$

13. **Application of equation** (25b). If the solution $u(x)$ of equation (25a) satisfies also (25b), the value of $L(u)$ may be eliminated between (25b) and (29) above. The result obtained is that

$$u(x) = \lambda\int_a^b\left\{K(x,\xi) + \frac{W(x)V(\xi)}{\Delta}\right\}u(\xi)d\xi, \tag{30}$$

and since the bracket in the integrand is, by (10), $K(x,\xi)$, this is precisely equation (5). This result we formulate as follows.

THEOREM 3. *Under the hypotheses of Theorem* 2 *every solution of the integro-differential system* (25) *is also a solution of the related normalized integral equation* (5).

The proof of the equivalence of system (25) and equation (5) is thus completed.

14. **Generalization.** In the deductions of this chapter we have been concerned hitherto with the particular system (25) derived in Chapter 3 from equation (5), and because of this we were able to identify as previously known functions the solutions of the various defining differential systems and integral equations. In any case, however, we see from equation (30) and the fact that $W(x)/\Delta$ and $K(x,\xi)$ there involved are solutions of the respective

equations (28), that the kernel of equation (30) is itself given as the solution of the integral equation

$$(31) \qquad K(x,\xi) = [G(x,\xi) + \Phi(x)V(\xi)] + \int_a^b F(x,t)K(t,\xi)dt.$$

It is clear from this that the method may be applied to the transformation of the general system of type (25), provided that the function $F(x, \xi)$ defined as above is possessed of a reciprocal. We may state therefore the following theorem.

THEOREM 4. *If an integro-differential system of type* (25) *satisfies the condition that the solution of its related differential system* (27) *is possessed of a reciprocal, then every solution of the integro-differential system solves also a related linear integral equation of the second kind.*

15. **The associated system.** Clearly all the methods employed above are applicable to the system $(\overline{25})$ as well as to (25). Hence we may conclude that system $(\overline{25})$ is equivalent to the equation $(\overline{5})$.

CHAPTER 5. THE EXISTENCE OF A SOLUTION OF THE INTEGRO-DIFFERENTIAL EQUATION FOR LARGE VALUES OF THE PARAMETER

16. **Matters of notation.** For the sake of simplicity in the formulations and deductions of this and subsequent chapters we shall make here certain conventions of notation. To begin with, the relation $|\lambda| > N$ shall be interpreted as an abbreviation of the statement "$|\lambda|$ sufficiently large." It is to be understood, therefore, that N does not necessarily mean the same constant in any one case as it does in any other. Further we reserve $B(x, \xi, \lambda)$ as a generic symbol for functions which for (x, ξ) in the region R and $|\lambda| > N$ are possessed of the following properties:

(a) $B(x, \xi, \lambda)$ is uniformly bounded, i.e.

$$|B(x, \xi, \lambda)| < B \text{ (a constant)};$$

(b) $B(x, \xi, \lambda)$ is integrable in ξ uniformly in x and λ, that is, if the interval (a, b) is subdivided by the points $\xi_0 = a < \xi_1 < \xi_2 < \cdots < \xi_n = b$ in any manner so that $|\xi_i - \xi_{i-1}|$ approaches zero as $n \to \infty$, and if $U_i(x, \lambda)$ and $L_i(x, \lambda)$ are respectively the upper and lower bounds of $B(x, \xi, \lambda)$ on the sub-interval (ξ_{i-1}, ξ_i), then when any constant $\epsilon > 0$ is prescribed there exists a numbe n_1 such that

$$\left| \sum_{i=1}^{n} [U_i(x,\lambda) - L_i(x,\lambda)](\xi_i - \xi_{i-1}) \right| < \epsilon$$

for $n \geqq n_1$.

Lastly σ and $i\tau$ will be used to designate the real and pure imaginary parts of λ, thus: $\lambda=\sigma+i\tau$, and a constant designated M shall be understood to be positive or zero.

17. Lemmas. For use in the subsequent deductions we begin by establishing the following lemmas.

LEMMA 1. *If* $a\leqq z\leqq b$ *and*

$$I(x,z,\lambda) = \int_a^z e^{\lambda(z-\xi)}B(x,\xi,\lambda)d\xi,$$

then

$$\lim_{\substack{|\lambda|\to\infty \\ \sigma\leqq M}} I = 0$$

uniformly in x and z and the argument of λ.

To establish this let the interval (a, b) be subdivided in the manner described under (b) above, the point z being taken as one of the points of subdivision, i.e. $\xi_k=z$. Then if $\bar{\xi}_i$ is any point of the interval (ξ_{i-1}, ξ_i) we have for ξ on this interval

$$| B(x,\xi,\lambda) - B(x,\bar{\xi}_i,\lambda) | < U_i(x,\lambda) - L_i(x,\lambda).$$

Let $I(x, z, \lambda)$ be written now in the form

$$I(x,z,\lambda) = \sum_{i=1}^{k} \left\{B(x,\bar{\xi}_i,\lambda) \int_{\xi_{i-1}}^{\xi_i} e^{\lambda(z-\xi)}d\xi \right.$$
$$\left. + \int_{\xi_{i-1}}^{\xi_i} e^{\lambda(z-\xi)}[B(x,\xi,\lambda) - B(x,\bar{\xi}_i,\lambda)]d\xi\right\}.$$

Then since

$$| e^{\lambda(z-a)} | \leqq e^{M(b-a)} \text{ for } \sigma \leqq M, \qquad a \leqq \xi \leqq z \leqq b,$$

we have

$$| I(x,z,\lambda) | \leqq \sum_{i=1}^{k} \left\{ B \cdot \left| \frac{e^{\lambda(z-\xi_{i-1})} - e^{\lambda(z-\xi_i)}}{\lambda} \right| \right.$$
$$\left. + e^{M(b-a)} [U_i(x,\lambda) - L_i(x,\lambda)] \int_{\xi_{i-1}}^{\xi_i} d\xi \right\}$$
$$\leqq \frac{2kBe^{M(b-a)}}{|\lambda|} + e^{M(b-a)} \sum_{i=1}^{k} [U_i(x,\lambda)$$
$$- L_i(x,\lambda)](\xi_i - \xi_{i-1}).$$

Now let $\epsilon>0$ be arbitrarily chosen. Then since $k\leqq n$ the final sum will, by hypothesis (b), be less than ϵ for $n\geqq n_1$, and we have

$$| I(x,z,\lambda) | \leqq e^{M(b-a)}\left[\frac{2n_1B}{|\lambda|}+\epsilon\right].$$

Since n_1 is fixed when ϵ is chosen, we may by making $|\lambda|$ sufficiently large reduce the right hand member of the inequality to less than $2e^{M(b-a)}\epsilon$. This proves the lemma.

LEMMA 2. *If*

$$I_2(x,z,\lambda) = \int_b^z e^{\lambda(z-\xi)}B(x,\xi,\lambda)d\xi$$

then

$$\lim_{\substack{|\lambda|\to\infty \\ \sigma\geqq -M}} I_2 = 0$$

uniformly in x and z and the argument of λ.

The proof of this lemma is in every way similar to that of Lemma 1.

18. Transformation of the integro-differential equation. We consider now the integro-differential equation

$$u'(x) - \lambda u(x) = L(u)\varphi(x) + \int_a^b f(x,\xi)u(\xi)d\xi + E(x,\lambda). \tag{32}$$

The symbol $E(x, \lambda)$ is to be considered generical, designating a uniformly bounded function which is analytic in λ for $|\lambda|>N$. The equation (32) reduces to equation (25a) in the special case $E(x, \lambda)\equiv 0$. For subsequent applications it is essential to consider the more general case.

If we consider equation (32) for the moment as a non-homogeneous equation of the type

$$u' + pu = q,$$

it follows from the theory of differential equations that it may be written in the form

$$u(x) = ce^{\lambda x} + \int_*^x e^{\lambda(x-t)}\left\{L(u)\varphi(t) + \int_a^b f(t,\xi)u(\xi)d\xi + E(t,\lambda)\right\}dt, \tag{32a}$$

the limit $*$ being any constant on the interval (a, b), and the coefficient c being constant with respect to x but otherwise arbitrary.

For purposes of orientation we shall proceed to deduce the consequences of the assumption that for such choices of the constants as will be made, equation (32a) is possessed of a solution. This deduction will be made

separately for the cases in which λ is confined to a left hand half-plane $\sigma \leqq M$, and in which λ is confined to a right hand half-plane $\sigma \geqq -M$.

19. **Heuristic deductions when** $\sigma \leqq M$. Let the arbitrary elements in (32a) be chosen in this case as follows:

$$* = a, \qquad c = c_1 e^{-\lambda a},$$

where c_1 is any absolute constant. The equation becomes, then,

$$\text{(32b)} \quad u(x) = c_1 e^{\lambda(x-a)} + \int_a^x e^{\lambda(x-t)} \left\{ L(u)\varphi(t) + \int_a^b f(t,\xi)u(\xi)d\xi + E(t,\lambda) \right\} dt .$$

Substituting the right hand member of this for $u(x)$ in the expression $L(u)$, we obtain, upon solving for $L(u)$,

$$\text{(33)} \quad L(u) = C^{-1} \left\{ L(c_1 e^{\lambda(x-a)}) + \nu \int_a^b e^{\lambda(b-t)} \left\{ E(t,\lambda) + \int_a^b f(t,\xi)u(\xi)d\xi \right\} dt \right\},$$

where

$$C = \left[1 - \nu \int_a^b e^{\lambda(b-t)} \varphi(t) dt \right].$$

This solution is possible if $|\lambda| > N$. For, since $\varphi(\xi)$ is a function of the type denoted by $B(x, \xi, \lambda)$ the integral in the expression for C decreases to zero as $|\lambda|$ increases, by Lemma 1. Hence $C \neq 0$ for $|\lambda| > N$.

With the value of $L(u)$ thus obtained equation (32b) takes the form

$$\text{(32c)} \qquad u(x) = \theta(x,\lambda) + \int_a^b \Omega(x,\xi,\lambda)u(\xi)d\xi,$$

where

$$\begin{aligned} \theta(x,\lambda) = {} & c_1 e^{\lambda(x-a)} + \int_a^x e^{\lambda(x-t)} E(t,\lambda) dt \\ & + C^{-1} \left\{ L(c_1 e^{\lambda(x-a)}) + \nu \int_a^b e^{\lambda(b-t)} E(t,\lambda) dt \right\} \cdot \int_a^x e^{\lambda(x-t)} \varphi(t) dt, \end{aligned}$$

and

$$\Omega(x,\xi,\lambda) = \int_a^x e^{\lambda(x-t)} f(t,\xi) dt + C^{-1}\nu \int_a^b e^{\lambda(b-t)} f(t,\xi) dt \cdot \int_a^x e^{\lambda(x-t)} \varphi(t) dt.$$

The assumed solution of equation (32b) must, therefore, satisfy also equation (32c).

20. **The existence of a solution of equation** (32) **for** $\sigma \leqq M$. Returning now to equation (32a) we may construct, by the respective formulas above,

the constant C, the functions $\theta(x, \lambda)$ and $\Omega(x, \xi, \lambda)$, the equation (32c), and the series

$$
(34)\qquad \theta(x,\lambda) + \int_a^b \Omega(x,\xi,\lambda)\theta(\xi,\lambda)d\xi + \int_a^b \Omega(x,\xi,\lambda)\int_a^b \Omega(\xi,\xi_1,\lambda)\theta(\xi_1,\lambda)d\xi_1 d\xi + \cdots .
$$

In the formulas for $\theta(x, \lambda)$ and $\Omega(x, \xi, \lambda)$ the integrals involving $E(t, \lambda)$ are uniformly bounded. The remaining integrals may be made arbitrarily small by taking $|\lambda| > N$, since $f(x, \xi)$, like $\varphi(\xi)$, is a function of the type $B(x, \xi, \lambda)$. Lastly $e^{\lambda(x-a)}$ is bounded, since $\sigma \leqq M$. Hence for $|\lambda| > N$

$$
|\theta(x,\lambda)| < A \qquad (\text{a constant}),
$$

$$
|\Omega(x,\xi,\lambda)| < \frac{1}{2(b-a)},
$$

and the terms of the series (34) are in absolute value less than the corresponding terms of the series

$$
A + \frac{A}{2} + \frac{A}{2^2} + \cdots .
$$

The series (34) converges, therefore, uniformly to a function numerically less than $2A$. Since the terms of (34) are continuous in x and analytic in λ it follows that (34) converges to a function $u(x, \lambda)$ which is uniformly bounded, continuous in x and analytic in λ. On the other hand by the classical theory of integral equations* this function $u(x, \lambda)$ is a solution of (32c).

We observe that if $E(x, \lambda) \equiv 0$, then the choice $c_1 = 0$ leads to $\theta(x, \lambda) \equiv 0$. The solution found is in this case $u(x, \lambda) \equiv 0$. We shall suppose, therefore, in proceeding, that c_1 has been so chosen that $|c_1| + |E(x, \lambda)| \not\equiv 0$.

Let the right hand member of (32c) be substituted now for $u(x)$ in the expression $L(u)$. The result is found to be precisely (33) above. The right hand member of (33) is contained, however, in the right hand member of (32c). The elimination of it between the two equations yields equation (32b), which is, therefore, satisfied by the function $u(x, \lambda)$ found. Lastly differentiating (32b) we obtain equation (32), which completes the proof that for $\sigma \leqq M$, $|\lambda| > N$, equation (32) is possessed of a solution which is uniformly bounded, continuous in x, and analytic in λ.

---

*Bôcher, loc. cit., p. 15.

21. The case $\sigma \geqq -M$. When λ is to be confined to a right hand half plane $\sigma \geqq -M$ the arbitrary elements in (32a) may be chosen as follows:

$$s = b, \qquad c = c_2 e^{-\lambda b},$$

where c_2 is any absolute constant. Reasoning precisely analogous to that followed in § 19 and § 20 serves, then, to establish the fact that for $\sigma \geqq -M$ and $|\lambda| > N$ equation (32) is also possessed of a solution with the characteristics noted above.

22. The associated equation. The discussion of the equation obtained from $(\overline{25\text{a}})$ by adding on the right a function $E(x, \lambda)$ may be carried through precisely like the discussion just concluded. In this way it is found that for λ confined to any half-plane $\sigma \leqq M$ or $\sigma \geqq -M$, and $|\lambda| > N$, such an equation also is possessed of a uniformly bounded solution which is analytic in λ and continuous in x.

Summarizing the results of this chapter we have obtained the following theorem.

THEOREM 5. *If the complex parameter λ is restricted to a region bounded by a line parallel to the axis of imaginaries and exterior to a circle sufficiently large and with center at $\lambda = 0$, then the integro-differential equation* (32) *admits of a solution which is uniformly bounded, continuous in x, and analytic in λ.*

CHAPTER 6. THE FORMAL SOLUTION OF THE INTEGRO-DIFFERENTIAL EQUATION

23. A lemma. In order to preclude interruption of the deductions about to be made we begin by establishing the following lemma. The symbol $H(x, \xi, \lambda)$ will be used to designate a function which is merely bounded uniformly for $|\lambda| > N$, and is integrable in x and ξ.

LEMMA 3. *If z_1 and z_2 are any points of the interval (a, b), and*

$$I(z_1, z_2, \xi, \lambda) = \int_{z_1}^{z_2} e^{\lambda t} H(\xi, t, \lambda)\, dt,$$

then the functional form of this integral is given by

$$I(z_1, z_2, \xi, \lambda) = e^{\lambda z_2} H(z_2, \xi, \lambda) + e^{\lambda z_1} H(z_1, \xi, \lambda).$$

This lemma follows almost immediately from the preceding ones. Thus the integral may be written in the form

$$I(z_1, z_2, \xi, \lambda) = e^{\lambda z_2} \int_s^{z_2} e^{-\lambda(z_2 - t)} H(\xi, t, \lambda)\, dt - e^{\lambda z_1} \int_s^{z_1} e^{-\lambda(z_1 - t)} H(\xi, t, \lambda)\, dt,$$

where * denotes any constant on the interval (a, b). If $-\sigma \leqq M$ we may choose $*=a$. Then the factors of the integrands on the right are all uniformly bounded. The same is therefore obviously true of the integrals themselves. On the other hand if $-\sigma \geqq -M$ we may choose $*=b$, with the result that the integrals are again uniformly bounded. This proves the lemma.

24. **Heuristic investigation of the functional form of a solution of (25a).** The existence of a solution of (25a) has been shown. We shall suppose now for purposes of orientation that such a solution may be obtained by the method of successive approximations, and in particular by the approximating scheme defined by the formula

$$u_i'(x) - \lambda u_i(x) = L(u_{i-1})\varphi(x) + \int_a^b f(x,\xi)u_{i-1}(\xi)d\xi. \tag{35}$$

This formula is in the form of a differential equation for $u_i(x)$ when $u_{i-1}(x)$ is known.

With the initial approximating function $u_0 \equiv 0$ we have as a possible choice $u_1 = e^{\lambda x}$. Substituting this in the right hand member of (35) and applying Lemma 3 we find that $u_2(x)$ is given by an equation of the type

$$u_2'(x) - \lambda u_2(x) = e^{\lambda b}H(x,\lambda) + e^{\lambda a}H(x,\lambda). \tag{35a}$$

The formula

$$u_2(x) = e^{\lambda x} + \int_*^x e^{\lambda(x-t)}\left\{e^{\lambda b}H(t,\lambda) + e^{\lambda a}H(t,\lambda)\right\}dt$$

yields a solution when the * is chosen arbitrarily for each term of the integrand. We obtain a particular solution $u_2(x)$, then, by choosing $*=b$ for the terms of which $e^{\lambda b}$ is a factor, and $*=a$ for the terms which contain e^{λ}. By Lemma (3), then

$$u_2(x) = e^{\lambda b}H(x,\lambda) + e^{\lambda x}H(x,\lambda) + e^{\lambda a}H(x,\lambda). \tag{36}$$

Upon the substitution of this as $u_{i-1}(x)$ in (35) and the application of Lemma 3 we find that $u_3(x)$ satisfies an equation of precisely the type (35a). Hence there exists a $u_3(x)$ of the same functional character as $u_2(x)$ given by (36). Obviously the same argument may be applied in the course of each successive substitution, and since we may consequently choose $u_n(x)$ for every n of the form of $u_2(x)$ above, the convergence of the procedure would lead to a solution of equation (35a) of the form (36). Guided by this result we proceed to substitute the form (36) into the equation with the purpose of determining as far as possible the functions $H(x, \lambda)$.

25. **The formal determination of coefficients.** The undetermined functions $H(x, \lambda)$ are bounded for $|\lambda| > N$. We shall proceed on the hypothesis that they are expansible in power series in $1/\lambda$, and shall write the assumed form of the solution of (25a)

$$u(x) = e^{\lambda x}[\gamma(x)] + e^{\lambda b}[\beta(x)] + e^{\lambda a}[\alpha(x)]. \tag{37}$$

In the notation used here

$$[\omega(x)] = \sum_{i=0}^{\infty} \frac{\omega_i(x)}{\lambda} \cdot$$

The result of the substitution of form (37) in the equation (25a) involves the integral

$$I = \int_a^b e^{\lambda\xi} f(x,\xi)[\gamma(\xi)]d\xi.$$

We shall consider this in two parts, namely

$$I_1 = \int_a^b e^{\lambda\xi} f(x,\xi) \sum_{k=0}^{n-2} \frac{\gamma_k(\xi)}{\lambda^k} d\xi,$$

and

$$I_2 = \int_a^b e^{\lambda\xi} f(x,\xi) \sum_{k=n-1}^{\infty} \frac{\gamma_k(\xi)}{\lambda^k} d\xi \cdot$$

Integrating the kth term of I_1 by parts $(n-2-k)$ times, we obtain the formula

$$I_1 = \sum_{k=0}^{n-3} \sum_{i=0}^{n-3-k} (-1)^{i+1} \frac{e^{\lambda b} f\gamma_k^i(x,b) - e^{\lambda x} Jf\gamma_k^i - e^{\lambda a} f\gamma_k^i(x,a)}{\lambda^{k+i+1}}$$
$$+ \frac{(-1)^{n-2}}{\lambda^{n-2}} \int_a^b e^{\lambda\xi} \sum_{k=0}^{n-2} (-1)^k f\gamma_k^{n-2-k}(x,\xi)d\xi.$$

The notation has been abbreviated here so that

$$f(x,\xi)\gamma_k(\xi) = f\gamma_k(x,\xi),$$

$$f\gamma_k(x,\xi)\Big]_{\xi=x}^{\xi=x+} = Jf\gamma_k,$$

$$\frac{\partial^i f\gamma_k(x,\xi)}{\partial\xi^i} = f\gamma_k^i(x,\xi).$$

The substitution of (37) in the equation yields therefore

$$
(38)\qquad
\begin{aligned}
&u'(x) - \lambda u(x) - L(u)\varphi(x) - \int_a^b f(x,\xi)u(\xi)d\xi \\
&= e^{\lambda x}\left\{[\gamma'(x)] + \sum_{k=0}^{n-3}\sum_{i=0}^{n-3-k}(-1)^i \frac{Jf_k^i}{\lambda^{k+i+1}}\right\} + e^{\lambda b}\left\{ \qquad \right\} \\
&\quad + e^{\lambda a}\left\{ \qquad \right\} + \frac{(-1)^{n-2}}{\lambda^{n-2}}\int_a^b e^{\lambda\xi}\sum_{k=0}^{n-2}(-1)^k f\gamma_k^{n-2-k}(x,\xi)d\xi \\
&\quad + \int_a^b e^{\lambda\xi} f(x,\xi)\sum_{k=n-1}^{\infty}\frac{\gamma_k(\xi)}{\lambda^k}\,d\xi = 0,
\end{aligned}
$$

where the coefficients of $e^{\lambda b}$ and $e^{\lambda a}$ are of the same general character as that of $e^{\lambda x}$.

Equating now to zero the coefficient of $e^{\lambda x}/\lambda^l$ for $l=0, 1, 2, \cdots, (n-2)$ respectively we find that $\gamma_0' \equiv 0$, whence we may choose $\gamma_0 = 1$, and that the functions $\gamma_1(x)$ $\gamma_2(x), \cdots, \gamma_{n-2}(x)$ can be successively determined each by means of a quadrature. Continuing in similar fashion and equating to zero the coefficients of $e^{\lambda b}/\lambda^l$ and $e^{\lambda a}/\lambda^l$ for $l=-1, 0, 1, \cdots, (n-3)$, respectively, we find likewise that

$$
(39)\qquad
\begin{aligned}
\beta_0 &= 0, \\
\beta_1(x) &= -\nu\varphi(x), \\
&\text{etc.}, \\
\alpha_0 &= 0, \\
\alpha_1(x) &= -\mu\varphi(x), \\
&\text{etc.},
\end{aligned}
$$

and that $\beta_2, \cdots, \beta_{n-2}, \alpha_2, \cdots, \alpha_{n-2}$ may also be successively evaluated. Hence we may construct the function $\bar{u}(x)$ given by the formula

$$
(40)\qquad
\begin{aligned}
\bar{u}(x) &= e^{\lambda x}\left\{1 + \frac{\gamma_1(x)}{\lambda} + \cdots + \frac{\gamma_{n-2}(x)}{\lambda^{n-2}}\right\} \\
&\quad + e^{\lambda b}\left\{\frac{\beta_1(x)}{\lambda} + \cdots + \frac{\beta_{n-2}(x)}{\lambda^{n-2}}\right\} \\
&\quad + e^{\lambda a}\left\{\frac{\alpha_1(x)}{\lambda} + \cdots + \frac{\alpha_{n-2}(x)}{\lambda^{n-2}}\right\}.
\end{aligned}
$$

26. **The equation satisfied by** $\bar{u}(x)$. The deductions above have been purely formal. A consideration of formulas involved shows, however, that

since $\varphi(x)$ is possessed of derivatives of order $(n-2)$, while $f(x, \xi)$ is possessed of partial derivatives of the same order, $\gamma_i(x)$ may be differentiated $(n-i)$ times, while $\beta_i(x)$ and $\alpha_i(x)$ are differentiable $(n-1-i)$ times. The function $\tilde{u}(x)$ determined is, therefore, differentiable and can be substituted into equation (25a).

The result of this substitution may be read directly from (38) above if the coefficients with subscripts greater than $(n-2)$ are taken now to be zero, and the symbols $[\gamma(x)]$ etc. are interpreted as representing the respective polynomials instead of the formal infinite series. The coefficients have been explicitly determined so that the terms in $e^{\lambda x}/\lambda^i$, $e^{\lambda b}/\lambda^{i-1}$, and $e^{\lambda a}/\lambda^{i-1}$ vanish for $i=0, 1, \cdots, (n-2)$. Since the integral I_2 of (38) does not now appear, there remain merely a term of the form $B(x)e^{\lambda b}/\lambda^{n-2}$, an analogous term in $e^{\lambda a}/\lambda^{n-2}$, and the integral

$$\frac{(-1)^{n-2}}{\lambda^{n-2}}\int_a^b e^{\lambda\xi} \sum_{k=0}^{n-2} (-1)^k f\gamma_k^{n-2-k}(x,\xi)d\xi.$$

Writing this integral in the form used in the proof of Lemma 3 we find readily that the substitution yields the result

$$\begin{aligned} &\tilde{u}'(x) - \lambda\tilde{u}(x) - L(\tilde{u})\varphi(x) - \int_a^b f(x,\xi)\tilde{u}(\xi)d\xi \\ (41) \qquad &= \frac{1}{\lambda^{n-2}}\left\{e^{\lambda b}B(x,\lambda) + e^{\lambda a}B(x,\lambda)\right\}, \end{aligned}$$

where, moreover, the functions $B(x, \lambda)$ are analytic in λ. We shall call the function $\tilde{u}(x)$ a formal solution to the order $(n-2)$ of equation (25a), since for such values of λ for which the exponential factors in (41) are bounded the equations (41) and (25a) agree to terms of degree $(n-2)$ in $(1/\lambda)$.

27. The associated equation. Equation $(\overline{25a})$ differs from equation (25a) in that each function involved is replaced by the corresponding function superscribed with a bar, and in that λ is replaced by $-\lambda$. With these alterations formula (40) above yields the formal solution of $(\overline{25a})$. Because of relation (26b), therefore, this formal solution is

$$\begin{aligned} &\bar{v}(x) = e^{-\lambda x}\left\{1 - \frac{\gamma_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\gamma}_{n-2}(x)}{\lambda^{n-2}}\right\} \\ (\overline{40}) \qquad &+ e^{-\lambda b}\left\{-\frac{\bar{\beta}_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\beta}_{n-2}(x)}{\lambda^{n-2}}\right\} \\ &+ e^{-\lambda a}\left\{-\frac{\bar{\alpha}_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\alpha}_{n-2}(x)}{\lambda^{n-2}}\right\}, \end{aligned}$$

the function $\gamma_1(x)$ occurring here being that which occurs also in formula (40).

CHAPTER 7. THE SOLUTION OF THE INTEGRO-DIFFERENTIAL EQUATION

28. **Relation of the formal solution to a true solution for $\sigma \leqq M$.** By the conclusions of § 20 equation (25a) admits, for $\sigma \leqq M$, $|\lambda| > N$, of a solution $u(x)$ which satisfies the relation (32b) with c_1 any non-vanishing constant, and $E \equiv 0$. Such a solution will not vanish for $x = a$, and hence upon being multiplied by a suitable function of λ alone, will satisfy the condition

$$u(a) = \bar{u}(a),$$

where $\bar{u}(x)$ is the formal solution (40). This multiplication by a function independent of x is permissible since the equation (25a) is homogeneous. Inasmuch as the solution of (32b) originally chosen is analytic in λ for $|\lambda| > N$, while the same is true of the formal solution $\bar{u}(x)$, the analyticity of the final form $u(x)$ is assured.

Consider now the function

$$w(x) = \lambda^{n-2} e^{-\lambda a}\{u(x) - \bar{u}(x)\}. \tag{42}$$

The bracket on the right obviously satisfies equation (41). Hence $w(x)$ is a solution of an equation

$$w'(x) - \lambda w(x) = L(w)\varphi(x) + \int_a^b f(x,\xi)w(\xi)d\xi + \{e^{\lambda(b-a)}B(x,\lambda) + B(x,\lambda)\},$$

the functions $B(x, \lambda)$ being analytic in λ. This equation is of the form (32) since for $\sigma \leqq M$ the bracket on the right is of the type $E(x, \lambda)$, and hence we may apply to it the deductions of § 19 and § 20. Since $w(a) = 0$ by (42), we find from the form (32b) that we have before us a case in which the coefficient c_1 of § 19 is zero. The function $\theta(x, \lambda)$ reduces, therefore, in this case to

$$\theta(x,\lambda) = \int_a^x e^{\lambda(x-t)}\{e^{\lambda(b-a)}B(t,\lambda) + B(t,\lambda)\}dt$$
$$+ C^{-1}\nu \int_a^b e^{\lambda(b-t)}\{e^{\lambda(b-a)}B(t,\lambda) + B(t,\lambda)\}dt \cdot \int_a^x e^{\lambda(x-t)}\varphi(t)dt,$$

and hence is, by Lemma 1, of the form $\epsilon(x, \lambda)$, where this symbol is used to denote a function which approaches zero uniformly in x as $|\lambda| \to \infty$. In consequence of this, as may be drawn from § 20, $w(x)$ is itself of the form $\epsilon(x, \lambda)$.

Solving (42), then, for $u(x)$ we obtain the result that there exists for $\sigma \leqq M$, $|\lambda| > N$, a solution of the form

$$u(x) = \bar{u}(x) + \frac{\epsilon(x,\lambda)e^{\lambda a}}{\lambda^{n-2}}.$$

This solution is moreover continuous in x and analytic in λ.

29. **The solution for $\sigma \geqq -M$.** The deductions of § 21 establish the existence of a solution of equation (25a) which for $\sigma \geqq -M$, $|\lambda| > N$, is analytic in λ, continuous in x, and non-vanishing at $x = b$. Because of the homogeneity of (25a) this solution may further be made to satisfy the condition

$$u(b) = \bar{u}(b).$$

We may conclude, then, in a manner entirely analogous to that of § 28, that the function

$$z(x) = \lambda^{n-2}e^{-\lambda b}\{u(x) - \bar{u}(x)\}$$

satisfies an equation of the type (32) and, since it vanishes at $x = b$, is of the form $\epsilon(x, \lambda)$. From this it follows further in the manner of § 28 that there exists a solution of equation (25a) which, for $|\lambda| > N$, $\sigma \geqq -M$, is analytic in λ and continuous in x, and which is of the form

$$u(x) = \bar{u}(x) + \frac{\epsilon(x,\lambda)e^{\lambda b}}{\lambda^{n-2}}.$$

Inasmuch as zero is a particular function of the type $\epsilon(x, \lambda)$ the results above may be summarized as follows.

THEOREM 6. *If the complex parameter λ is restricted to a region bounded by a line parallel to the axis of imaginaries and exterior to a circle sufficiently large and with center at $\lambda = 0$, then the integro-differential equation* (25a) *admits of a solution of the form*

$$\begin{aligned} u(x) &= e^{\lambda x}\left\{1 + \frac{\gamma_1(x)}{\lambda} + \cdots + \frac{\gamma_{n-2}(x)}{\lambda^{n-2}}\right\} \\ &\quad + e^{\lambda b}\left\{\frac{\beta_1(x)}{\lambda} + \cdots + \frac{\beta_{n-2}(x)}{\lambda^{n-2}} + \frac{\epsilon(x,\lambda)}{\lambda^{n-2}}\right\} \qquad (43) \\ &\quad + e^{\lambda a}\left\{\frac{\alpha_1(x)}{\lambda} + \cdots + \frac{\alpha_{n-2}(x)}{\lambda^{n-2}} + \frac{\epsilon(x,\lambda)}{\lambda^{n-2}}\right\}. \end{aligned}$$

This solution is analytic in λ and continuous in x.

30. **The associated equation.** Clearly the analysis applied above to the equation (25a) and its formal solution (40), may be applied equally well to the associated equation $(\overline{25\text{a}})$ and its formal solution $(\overline{40})$. In this manner it is found that when λ is confined to any right or left-hand half-plane and $|\lambda| > N$, the equation $(\overline{25\text{a}})$ admits of a solution $v(x)$ which is continuous in x, analytic in λ, and of the form

$$
\begin{aligned}
v(x) &= e^{-\lambda x}\left\{1 - \frac{\gamma_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\gamma}_{n-2}(x)}{\lambda^{n-2}}\right\} \\
&+ e^{-\lambda b}\left\{-\frac{\bar{\beta}_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\beta}_{n-2}(x)}{\lambda^{n-2}} + \frac{\epsilon(x,\lambda)}{\lambda^{n-2}}\right\} \qquad (\overline{43}) \\
&+ e^{-\lambda a}\left\{-\frac{\bar{\alpha}_1(x)}{\lambda} + \cdots + (-1)^{n-2}\frac{\bar{\alpha}_{n-2}(x)}{\lambda^{n-2}} + \frac{\epsilon(x,\lambda)}{\lambda^{n-2}}\right\}.
\end{aligned}
$$

We remark lastly that since the coefficients and functions involved in equation (25a) and $(\overline{25\text{a}})$ are all real, conjugate imaginary values of λ will correspond to the conjugate imaginary solutions $u(x)$ or $v(x)$. Thus if we denote by λ_c, $u_c(x)$, and $v_c(x)$ the conjugates of λ, $u(x)$, and $v(x)$ respectively,

$$
\begin{aligned}
u(x,\lambda_c) &= u_c(x,\lambda), \\
v(x,\lambda_c) &= v_c(x,\lambda).
\end{aligned}
$$

Chapter 8. The characteristic values

31. **The characteristic equation.** The immediately preceding sections have been concerned with the form of a solution of equation (25a). By Chapters 3 and 4 such a solution solves the given normalized integral equation if and only if equation (25b) is also satisfied. This imposes a restriction on the choice of λ which we proceed to consider.

Substituting for the factor $\lambda u(t)$ in (25b) its value as the second term of (25a) we obtain the equation

$$
L(u) = \int_a^b V(t)\left\{u'(t) - L(u)\varphi(t) - \int_a^b f(t,\xi)u(\xi)d\xi\right\}dt,
$$

which, if we integrate by parts the term involving $u'(t)$, takes the form

$$
k_1 u(a) - k_2 u(b) + \int_a^b \Delta \cdot \bar{\varphi}(\xi)u(\xi)d\xi.
$$

In this, $\Delta\cdot\bar{\varphi}(\xi)$ is given by formula (26a), and

$$\text{(44a)}\qquad \begin{aligned} k_1 &= V(a) + \mu\left\{1 + \int_a^b V(\xi)\varphi(\xi)d\xi\right\}, \\ k_2 &= V(b) - \nu\left\{1 + \int_a^b V(\xi)\varphi(\xi)d\xi\right\}. \end{aligned}$$

It is convenient to set

$$\text{(44b)}\qquad \begin{aligned} \Omega_1(\xi) &= \frac{-\Delta\cdot\bar{\varphi}(\xi)}{2k_1}, \\ \Omega_2(\xi) &= \frac{\Delta\cdot\bar{\varphi}(\xi)}{2k_2}, \end{aligned}$$

and to write the equation in the form

$$\text{(45)}\qquad k_1\left[u(a) - \int_a^b \Omega_1(\xi)u(\xi)d\xi\right] - k_2\left[u(b) - \int_a^b \Omega_2(\xi)u(\xi)d\xi\right] = 0.$$

We shall call this equation the characteristic equation. Its roots, if such exist, are values of λ for which the solution of the integro-differential equation satisfies also the given integral equation. We shall call these roots characteristic values and denote them by λ_i.

We observe that since the coefficients of equation (45) are all real, the equation is essentially unchanged if $u(x)$ is replaced by its conjugate imaginary $u_c(x)$. From this it follows, by § 30, that the characteristic values occur in conjugate pairs.

Now for λ in any right or left hand half-plane we have by Theorem 6 a solution

$$u(x) = e^{\lambda x} + e^{\lambda b}\epsilon(x,\lambda) + e^{\lambda a}\epsilon(x,\lambda),$$

which is analytic in λ for $|\lambda| > N$. The substitution of this in (45) yields for the characteristic equation the form

$$\text{(46)}\qquad [k_1 + \epsilon(\lambda)]e^{\lambda a} - [k_2 + \epsilon(\lambda)]e^{\lambda b} = 0,$$

in which the functions $\epsilon(\lambda)$ are analytic for $|\lambda| > N$.

32. **A further condition on the given equation.** It is essential to the procedure which is to be followed in solving equation (46) that neither of the coefficients k_1 and k_2 shall vanish. Since k_1 and k_2 are seen from the formulas to be rational functions of μ and ν we may choose μ and ν subject to previous restrictions so that $k_1k_2 \neq 0$, provided only that neither k_1 nor k_2 vanishes

identically in μ and ν. We shall suppose then that the given integral equation is such that

(vi) $$k_1(\mu,\nu) \not\equiv 0, \qquad k_2(\mu,\nu) \not\equiv 0.$$

From the formulas for k_1 and k_2 it follows immediately that

(47) $$\nu k_1(\mu,\nu) + \mu k_2(\mu,\nu) = \Delta(\mu,\nu).$$

It is a clear from this that a sufficient condition for (vi) is that

$$\Delta(0,\nu) \neq 0, \qquad D(0,\nu) \neq 0,$$

and

$$\Delta(\mu,0) \neq 0, \qquad D(\mu,0) \neq 0.$$

For if these conditions are fulfilled k_1 and k_2 are defined for the values $(\mu, 0)$ and $(0, \nu)$ and if either one vanished identically relation (47) would entail a contradiction for $\mu=0$ or for $\nu=0$.

The restriction (vi) is an essential one, for we have in the Volterra equations, i.e. those in which $K(x, \xi)\equiv 0$ in one of the regions R_1 and R_2, an example of equations which may satisfy all previous conditions, but to which the subsequent results do not apply. To show that these equations are ruled out by condition (vi) we may proceed as follows.

Let us suppose $K(x, \xi)\equiv 0$ for $\xi > x$. Then $W(a)=\nu$ and $\nu\neq 0$ because of (8). We recall that $E(x, \xi)=K_\xi(x, \xi)$, and hence that Fredholm's identity (15) yields the formula

$$K_\xi(a,\xi) + F(a,\xi) = \int_a^b K_t(a,t)F(t,\xi)dt.$$

Upon an integration by parts the right hand member of this assumes the form

$$K(a,b)F(b,\xi) - K(a,a+)F(a,\xi) - \int_a^b K(a,t)F_t(t,\xi)dt,$$

of which the first two terms reduce because of (17b) and (12) to $F(a, \xi)$. Hence we may write

$$K_\xi(a,\xi) = -\int_a^b K(a,t)F_t(t,\xi)dt,$$

or, upon substituting from relation (10),

$$\int_a^b \frac{V(t)F_t(t,\xi)dt}{\Delta} = \frac{1}{W(a)}\left[K_\xi(a,\xi) + \int_a^b F_t(t,\xi)K(a,t)dt\right] - \frac{V'(\xi)}{\Delta}.$$

Multiplying this by $W(\xi)$ and integrating we obtain the form

$$\int_a^b V(t)\int_a^b F_t(t,\xi)\frac{W(\xi)}{\Delta}d\xi dt = \frac{1}{W(a)}\int_a^b W(\xi)\left[K_\xi(a,\xi)+\int_a^b F_t(t,\xi)K(a,t)dt\right]d\xi$$

$$-\frac{1}{\Delta}[W(b)V(b)-W(a)V(a)]+\int_a^b V(\xi)\frac{W'(\xi)}{\Delta}d\xi.$$

With the use of formulas (11) and a rearrangement of terms this leads finally to the relation

$$\int_a^b V(t)\varphi(t)dt = \frac{V(b)-W(a)}{\nu}$$

$$-\frac{1}{W(a)}\int_a^b W(\xi)\left[K_\xi(a,\xi)+\int_a^b F_t(t,\xi)K(a,t)dt\right]d\xi.$$

From this form it is readily found, now, that since $K(x,\xi)\equiv 0$ for $\xi > x$, the second term on the right vanishes, with the result that

$$k_1 = \mu + \nu K(b,a), \qquad k_2 \equiv 0.$$

33. **The solution of the characteristic equation.** We consider to begin with the related, simpler equation

$$k_1 e^{\rho a} - k_2 e^{\rho b} = 0. \tag{48}$$

Because of (vi) we may suppose μ and ν so chosen that $k_1\neq 0$, $k_2\neq 0$, and the equation is easily solvable. Thus $e^{\rho(b-a)} = k_1/k_2$, and by taking logarithms the roots are found to be

$$\rho_m = \frac{1}{b-a}\left\{2m\pi i + \log\frac{k_1}{k_2}\right\} \qquad (m = 0, \pm 1, \pm 2, \cdots). \tag{49}$$

Solving the equation (46) now formally in the same manner we find

$$\lambda_m = \frac{1}{b-a}\left\{2m\pi i + \log\frac{k_1+\epsilon(\lambda_m)}{k_2+\epsilon(\lambda_m)}\right\}.$$

Then inasmuch as in functional form

$$\log\frac{k_1+\epsilon(\lambda_m)}{k_2+\epsilon(\lambda_m)} = \log\frac{k_1}{k_2}+\epsilon(\lambda_m),$$

we have

$$\lambda_m = \rho_m + \epsilon(\lambda_m). \tag{46a}$$

Consider now the function

$$\theta(\lambda) = \lambda - [\rho_m + \epsilon(\lambda)],$$

the function ϵ being that of (46a). The function $\theta(\lambda)$ is analytic, and if $\delta>0$ be chosen arbitrarily but sufficiently small it will follow that $|\epsilon(\lambda)|<\delta$ for $|\lambda|>N$. If, therefore, a circle C_m with radius δ be drawn about each of the points ρ_m, then for $|m|$ sufficiently large the point $\rho_m+\epsilon(\lambda)$ will lie within the corresponding circle C_m whenever λ is a point on the circumference. As λ describes this circumference, therefore, the argument of $\theta(\lambda)$ increases by 2π, from which it follows that $\theta(\lambda)=0$ has precisely one root within C_m. This means if we denote this root by λ_m that the points λ_m are represented asymptotically by the points ρ_m, namely

$$\lambda_m = \rho_m + \epsilon_m, \tag{50}$$

where $\epsilon_1, \epsilon_2, \cdots, \epsilon_m, \cdots$ is a sequence of constants such that $\lim_{m\to\infty} \epsilon_m=0$. Moreover since the characteristic values occur in conjugate pairs, while ρ_m and ρ_{-m} are conjugate,* we conclude that

$$\lambda_{-m}=\bar{\lambda}_m.$$

From the significance of ρ_m for the given integral equation, and from the fact that by (49) ρ_m involves the ratio k_1/k_2, it follows that this ratio is determined by the integral equation alone and so must be independent† of the parameters μ and ν.

34. **A more precise formula for λ_m.** In deducing the formula (50) we assumed only that for the given integral equation $n\geqq 2$. If we have at our disposal the existence of derivatives of higher order a formula more precise than (50) may be deduced. We shall assume, now, for the following deductions that for the given integral equation

$$n \geqq 4. \tag{vii}$$

The assumption (vii) insures the existence of a solution $u(x)$ of equation (25a) which is of the form given by (43) for $n=4$. Substituting this solution in (45) and integrating by parts the integrals involving $e^{\lambda\xi}\Omega(\xi)$ we obtain for the characteristic equation the form

$$e^{\lambda a}\left\{k_1 + \frac{k_{12}}{\lambda} + \frac{k_{13}+\epsilon(\lambda)}{\lambda^2}\right\} - e^{\lambda b}\left\{k_2 + \frac{k_{22}}{\lambda} + \frac{k_{23}+\epsilon(\lambda)}{\lambda^2}\right\} = 0,$$

* It is assumed at this point that k_1/k_2 is positive. If it is negative we may write $\log(k_1/k_2) = \pi i + \log|k_1/k_2|$, in which case it is evident that ρ_m and ρ_{-m-1} and hence also λ_m and λ_{-m-1} are conjugate. To avoid unnecessary complications we shall continue only with the case $k_1/k_2>0$, the modifications for $k_1/k_2<0$ being reasonably obvious.

† The example of the appendix illustrates this.

where the coefficients k are constants and in particular

$$(51)\qquad \begin{aligned} k_{12} &= k_1\left\{\gamma_1(a) + \alpha_1(a) + 2\Omega_1(a) - 2\int_a^b \Omega_1(\xi)\alpha_1(\xi)d\xi\right\} - k_2\alpha_1(b), \\ k_{22} &= k_2\left\{\gamma_1(b) + \beta_1(b) - 2\Omega_2(b) - 2\int_a^b \Omega_2(\xi)\beta_1(\xi)d\xi\right\} - k_1\beta_1(a). \end{aligned}$$

Solving this equation formally in the fashion of § 33 we obtain

$$\lambda_m = \rho_m + \frac{1}{b-a}\log\frac{1 + \dfrac{k_{12}}{k_1\lambda_m} + \dfrac{k_{13} + \epsilon(\lambda_m)}{k_1\lambda_m^2}}{1 + \dfrac{k_{22}}{k_2\lambda_m} + \dfrac{k_{23} + \epsilon(\lambda_m)}{k_2\lambda_m^2}},$$

or, upon expanding the logarithm as we may, because of (50), for $|m|$ sufficiently large, the form

$$\lambda_m = \rho_m + \frac{1}{b-a}\left[\frac{\dfrac{k_{12}}{k_1} - \dfrac{k_{22}}{k_2}}{\lambda_m}\right] + \frac{A' + \epsilon(\lambda_m)}{\lambda_m^2},$$

A' representing a constant independent of m. From the formulas (46a) and (49) we find readily, however, that

$$\frac{1}{\lambda_m} = \frac{b-a}{2m\pi i} + \frac{A'' + \epsilon(m)}{m^2},$$

A'' being a constant. If this is substituted in the right hand member of the preceding expression we obtain a result which is expressed in the second part of the following theorem.

THEOREM 7. *If the kernel of the normalized integral equation* (5) *satisfies the conditions* (iii), (iv), (v) *and* (vi) *then there exist infinitely many values of the parameter* λ *for which the integral equation admits of a solution. These characteristic values are of the form*

$$\lambda_m = \frac{1}{b-a}\left\{2m\pi i + \log\frac{k_1}{k_2}\right\} + \epsilon_m \quad (m = 0, \pm 1, \pm 2, \cdots),$$

where

$$\lim_{|m|\to\infty} \epsilon_m = 0.$$

If the kernel of the integral equation satisfies the conditions (iv), (v), (vi) *and* (vii), *these characteristic values may be determined more precisely by the formula*

$$(50a)\qquad \lambda_m = \frac{1}{b-a}\left\{2m\pi i + \log\frac{k_1}{k_2}\right\} + \frac{(k_{12}/k_1) - (k_{22}/k_2)}{2m\pi i} + \frac{A + \epsilon_m}{m^2}.$$

The various constants k are given by formulas (44a) *and* (51), *and A also is a constant independent of m.*

Obviously if a larger number of derivatives of the kernel of the given equation are at hand the method applied above is applicable to the deduction of a formula still more precise than (50a).

35. **The associated equation.** It is not necessary in this case to repeat the discussion for the system $(\overline{25})$. By the classical theory of integral equations the associated equations (5) and $(\bar{5})$ have the same characteristic values. Hence formula (50a) serves both.

Chapter 9. The characteristic functions

36. **The solutions of the integral equation.** The characteristic values λ_m by their very determination are such that the functions $u(x, \lambda_m)$ satisfy not merely equation (25a) but the entire system (25). As such these functions are also solutions of the given normalized integral equation. For brevity we shall denote them by $u_m(x)$, and shall call them characteristic functions.

The form of the characteristic function $u_m(x)$ is obviously obtained by substituting the value $\lambda=\lambda_m$ in the form $u(x)$ given by formula (43). Dividing this formula by $e^{\lambda a}$ and observing that for $|m|$ sufficiently large $e^{\lambda_m(x-a)}$ is uniformly bounded, we obtain because of (vii) the formula

$$u_m(x) = e^{\lambda_m(x-a)}\left\{1+\frac{\gamma_1(x)}{\lambda_m}\right\}+e^{\lambda_m(b-a)}\frac{\beta_1(x)}{\lambda_m}+\frac{\alpha_1(x)}{\lambda_m}+\frac{H(x,\lambda_m)}{\lambda_m^2}.$$

Since by (50a)

$$\begin{aligned} e^{\lambda_m(x-a)} &= e^{\rho_m(x-a)+[(1/2m\pi i)((k_{12}/k_1)-(k_{22}/k_2))+H(m)/m^2](x-a)} \\ &= e^{\rho_m(x-a)}\left[1+\frac{x-a}{2m\pi i}\left(\frac{k_{12}}{k_1}-\frac{k_{22}}{k_2}\right)+\frac{H(x,m)}{m^2}\right], \end{aligned}$$

we find in terms of ρ_m and m

$$u_m(x) = e^{\rho_m(x-a)}\left\{1+\frac{Q_1(x)}{m}\right\}+\frac{Q_2(x)}{m}+\frac{H(x,m)}{m^2}, \tag{52}$$

where

$$Q_1(x) = \frac{\gamma_1(x)(b-a)+\big((k_{12}/k_1)-(k_{22}/k_2)\big)(x-a)}{2\pi i} \tag{53a}$$

and

$$Q_2(x) = \frac{b-a}{2\pi i}\left\{e^{\rho_m(b-a)}\beta_1(x)+\alpha_1(x)\right\}.$$

By (48), however,

$$e^{\rho_m(b-a)} = \frac{k_1}{k_2}\,.$$

Hence the function $Q_2(x)$ is independent of m and substituting for $\beta_1(x)$ and $\alpha_1(x)$ their values from formulas (39) we find that

$$Q_2(x) = \frac{-(b-a)}{2\pi i}\left[\mu + \frac{k_1}{k_2}\nu\right]\varphi(x). \tag{53b}$$

Deductions precisely similar to those above based on the function $(\overline{43})$ yield the characteristic functions of equation $(\overline{5})$. We have first

$$v_m(x) = e^{-\lambda_m(x-a)}\left\{1 - \frac{\gamma_1(x)}{\lambda_m}\right\} - e^{-\lambda_m(b-a)}\frac{\bar\beta_1(x)}{\lambda_m} - \frac{\bar\alpha_1(x)}{\lambda_m} + \frac{H(x,\lambda_m)}{\lambda_m{}^2}$$

from which in the manner above we obtain

$$v_m(x) = e^{-\rho_m(x-a)}\left\{1 - \frac{Q_1(x)}{m}\right\} + \frac{Q_3(x)}{m} + \frac{H(x,m)}{m^2}\,. \tag{$\overline{52}$}$$

In this $Q_1(x)$ is the function occurring also in (52), and

$$Q_3(x) = \frac{b-a}{2\pi i}\left[\nu + \mu\frac{k_2}{k_1}\right]\bar\varphi(x). \tag{$\overline{53\text{b}}$}$$

37. **The normalized set.** By the classical theory of integral equations the set of functions $u_m(x)$, $v_m(x)$ is biorthogonal, that is,

$$\int_a^b u_m(x)v_p(x)dx = 0, \qquad m \neq p.$$

The set of functions is said to be normalized if further

$$\int_a^b u_m(x)v_m(x)dx = 1.$$

Forming the product $u_m(x)\,v_m(x)$ from formulas (52) and $(\overline{52})$ we obtain

$$u_m(x)v_m(x) = 1 + \frac{Q_2(x)e^{-\rho_m(x-a)} + Q_3(x)e^{\rho_m(x-a)}}{m} + \frac{H(x,m)}{m^2}\,.$$

Then since $|\rho_m|$ is of the order m, while $Q_2(x)$ and $Q_3(x)$ are possessed of a derivative, we obtain as the result of an integration by parts

$$\int_a^b \frac{Q_i(x)e^{\pm\rho_m(x-a)}dx}{m} = \frac{H(m)}{m^2} \qquad (i = 2,3).$$

Accordingly

$$\int_a^b u_m(x)v_m(x)dx = (b-a) + \frac{H(m)}{m^2}\cdot$$

We may therefore normalize the solutions given by (52) and ($\overline{52}$) by multiplying them by a suitable factor of the form

$$\frac{1}{\sqrt{b-a}} + \frac{H(m)}{m^2}\cdot$$

Thus we obtain the following theorem.

THEOREM 8. *If the kernel of the normalized integral equation* (5) *satisfies the conditions* (iv), (v), (vi), *and* (vii), *and the kernel of the associated integral equation satisfies condition* ($\overline{\text{v}}$), *then the normalized characteristic solutions of these equations are given respectively, for* $|m|$ *sufficiently large, by the formulas*

$$(54)\qquad \begin{aligned} u_m(x) &= \frac{1}{\sqrt{b-a}}\left[e^{\rho_m(x-a)}\left\{1+\frac{Q_1(x)}{m}\right\}+\frac{Q_2(x)}{m}+\frac{H(x,m)}{m^2}\right],\\ v_m(x) &= \frac{1}{\sqrt{b-a}}\left[e^{-\rho_m(x-a)}\left\{1-\frac{Q_1(x)}{m}\right\}+\frac{Q_3(x)}{m}+\frac{H(x,m)}{m^2}\right].\end{aligned}$$

Finally we remark that by § 30 and § 33 the functions $u_m(x)$ and $u_{-m}(x)$, as well as the functions $v_m(x)$ and $v_{-m}(x)$, are conjugate imaginary.

CHAPTER 10. THE CLOSURE OF THE SET OF CHARACTERISTIC FUNCTIONS

38. **Birkhoff's theorem.** It has been shown in the preceding chapters that for $|m|$ sufficiently large the characteristic values of the given integral equation are simple, and that the corresponding biorthogonal set of characteristic functions $u_m(x)$, $v_m(x)$ may be normalized. We shall now make the further assumption:

(viii) All the characteristic values of the given integral equation are simple.

Then the entire set of solutions $u_m(x)$, $v_m(x)$ may be normalized. We wish to show that the set is also closed. To this end we shall employ the following theorem which is a generalization of a theorem given by Birkhoff* for orthogonal sets.

---

* Birkhoff, Proceedings of the National Academy of Sciences, vol. 3 (1917), p. 656.

THEOREM. *If $w_m(x)$, $z_m(x)$, $m=0, \pm 1, \pm 2, \cdots$, is a closed normalized biorthogonal set of functions, and if $\bar{w}_m(x)$, $\bar{z}_m(x)$ is a second normalized biorthogonal set, then this second set is also closed provided* (a) *the series*

$$[w_0(x) - \bar{w}_0(x)]z_0(y) + \sum_{m=1}^{\infty} \{ [w_m(x) - \bar{w}_m(x)]z_m(y) + [w_{-m}(x) - \bar{w}_{-m}(x)]z_{-m}(y) \}$$

converges to a function $H(x, y)$ less than $1/2(b-a)$ in numerical magnitude, and (b) *the convergence is such that the series multiplied through by an arbitrary continuous function $f(x)$ can be integrated term by term with respect to x, and yields a series which converges uniformly on any closed sub-interval of $a<y<b$, and of which the sum of any number of terms remains uniformly bounded for $a \leqq y \leqq b$.*

The proof, except for minor modifications, is that given by Birkhoff. Thus if the set $\bar{w}_m(x)$, $\bar{z}_m(x)$ is not closed, there exists a function $f(x) \not\equiv 0$ which is continuous on (a, b) and such that

$$\int_a^b f(x)\bar{w}_m(x)dx = 0 \quad \text{for all } m.$$

In this event if we multiply through the equation of definition for $H(x, y)$ by this $f(x)$ and integrate term by term as we may by hypothesis, we find

$$\int_a^b H(x,y)f(x)dx = \int_a^b f(x)w_0(x)z_0(y)dx + \sum_{m=1}^{\infty} \int_a^b f(x)[w_m(x)z_m(y) + w_{-m}(x)z_{-m}(y)]dx,$$

where the series on the right hand side converges under the hypothesis to a bounded function which is continuous for $a<y<b$. This function on the open interval is precisely $f(y)$. In fact the difference

$$f(y) - \left\{ \int_a^b f(x)w_0(x)z_0(y)dx + \sum_{m=1}^{\infty} \int_a^b f(x)[w_m(x)z_m(y) + w_{-m}(x)z_{-m}(y)]dx \right\}$$

is a function $\varphi(y)$ bounded on (a, b), continuous for $a<y<b$, and such that

$$\int_a^b \varphi(y)w_l(y)dy = 0 \quad \text{for all } l.$$

This latter property is readily drawn from the fact that the series involved may be integrated term by term, while the set $w_m(y)$, $z_m(y)$ is normalized

biorthogonal. But the set $w_m(y)$, $z_m(y)$ is closed by hypothesis. Hence we infer that $\varphi(y) \equiv 0$, $a < y < b$, and that the right hand member of the preceding equation has the value $f(y)$. That equation may now be written

$$\int_a^b H(x,y)f(x)dx = f(y), \qquad a < y < b.$$

If the upper bound of $|f(x)|$ on (a, b) is $F \neq 0$ we obtain from this since $|H(x, y)| < 1/2(b-a)$ the relation

$$|f(y)| < \tfrac{1}{2}F.$$

But this is impossible since the left hand member either takes on the value F or comes arbitrarily near to it. The set $\bar{w}_m(x)$, $\bar{z}_m(x)$ is therefore closed also.

39. **The integro-differential system with an additional parameter.** We consider now the system

$$(55)\quad \begin{aligned} &\text{(a)} \qquad u'(x) - \lambda u(x) = \eta\left[L(u)\varphi(x) + \int_a^b f(x,\xi)u(\xi)d\xi\right], \\ &\text{(b)} \qquad L(u) = \eta\lambda\int_a^b V(\xi)u(\xi)d\xi. \end{aligned}$$

For $\eta = 1$ this is simply the system (25) and defines the set of functions $u_m(x, 1) \equiv u_m(x)$. For $\eta = 0$ it is an ordinary differential system and defines a set $u_m(x, 0)$. For both these values of η there exists an associated or adjoint system with its corresponding set of solutions $v_m(x, 1)$ or $v_m(x, 0)$ which is biorthogonal to the respective set above. It is our purpose to show that for other suitable values of η in the circle $C : |\eta| \leqq 1$ of the complex η plane the system (55) also defines a set of functions $u_m(x, \eta)$ and that there exists a biorthogonal normal set $v_m(x, \eta)$. Subsequently it will be shown that the repeated application of Birkhoff's theorem makes it possible to conclude the closure of the set $u_m(x, 1)$, $v_m(x, 1)$ from the known closure of the set $u_m(x, 0)$, $v_m(x, 0)$.

The system (55) will be equivalent to an integral equation of the form (5) provided the hypotheses of Chapter 4 as applied to system (55) are met. The extension of the deductions of Chapter 4 to system (55) is made formally by replacing respectively $\varphi(x)$, $f(x, \xi)$, and $V(\xi)$, by $\eta\varphi(x)$, $\eta f(x, \xi)$ and $\eta V(\xi)$. This is readily found to result in replacing $F(x, \xi)$ by $\eta F(x, \xi)$, and $\Phi(x)$ by $\Phi(x, \eta) = \frac{1}{2}(1-\eta) + \eta\Phi(x)$. The essential hypothesis to be met is that the function $\eta F(x, \xi)$ be possessed of a reciprocal function $E(x, \xi, \eta)$. By the classical theory of integral equations this reciprocal is a meromorphic function

of η. Hence its poles within the circle C are finite in number and may be denoted by η_i. Then for η in any closed sub-region of C, call it C', which excludes these points η_i, the function $E(x, \xi, \eta)$ is analytic, and by Theorem 4 the system (55) is equivalent to an integral equation

$$u(x,\eta) = \lambda \int_a^b K(x,\xi,\eta)u(\xi,\eta)d\xi. \tag{56}$$

The kernel $K(x, \xi, \eta)$, since it satisfies the relation (31) with the formal modifications noted above, is readily found to be continuous in x and η, and in fact analytic in η in the region C'.

Consider now the associated equation

$$v(x,\eta) = \lambda \int_a^b K(\xi,x,\eta)v(\xi,\eta)d\xi, \tag{$\overline{56}$}$$

and the deductions of Chapter 3 as applied to it. From the analytic character of $K(\xi, x, \eta)$ and its relation to the kernel of (56), the function $\overline{E}(x, \xi, \eta)$ of § 11 is readily found to be analytic in η. From the classical theory of integral equations we draw again the fact that $\overline{E}(x, \xi, \eta)$ possesses a reciprocal $\overline{F}(x, \xi, \eta)$ which is meromorphic in η. If we add to the set of points η_i those poles of $\overline{F}(x, \xi, \eta)$ which lie within C' and denote by C'' any closed sub-region of C' which excludes also these values η_i, it follows that $\overline{F}(x, \xi, \eta)$ exists and is analytic in C''. By Chapter 3 it follows then that the equation ($\overline{56}$) is also equivalent to an integro-differential system which is the associated system of (55).

It is essential to observe that in obtaining this result we have not excluded from C'' either $\eta=1$ or $\eta=0$. The case $\eta=1$ is covered by the hypotheses of the preceding chapters. For $\eta=0$ the system (55) is equivalent to the integral equation

$$u(x,0) = \lambda \int_a^b G(x,\xi)u(\xi,0)d\xi,$$

and the associated equation

$$v(x,0) = \lambda \int_a^b G(\xi,x)v(\xi,0)d\xi$$

is equivalent to the system

$$v'(x,0) + \lambda v(x,0) = 0,$$
$$\overline{L}(v) = 0.$$

The extension of the deductions of Chapters 5, 6, and 7 to the system (55) and its associated system may be made now with the formal modifica-

tions noted. For the deductions of Chapters 8 and 9 the condition (vi), which in the present case takes the form

$$k_1(\eta) = \eta V(a) + \mu\left[1 + \eta^2\int_a^b V(\xi)\varphi(\xi)d\xi\right] \neq 0,$$

$$k_2(\eta) = \eta V(b) - \nu\left[1 + \eta^2\int_a^b V(\xi)\varphi(\xi)d\xi\right] \neq 0,$$

must be met. For $\eta=1$ this is assured by previous hypotheses. Hence k_1 and k_2 as functions of η are not identically zero, and there are at most four values of η for which either k_1 or k_2 will vanish. Let such of these values as lie in C'' be added to the set of points η_i and denote by C''' any closed subregion of C'' which excludes also these points. This does not exclude $\eta=1$, and if we suppose, as we may and shall, that $\mu\neq 0$, $\nu\neq 0$, it does not exclude $\eta=0$. For η in the region C''', then, the biorthogonal set of functions $u_m(x, \eta)$, $v_m(x, \eta)$ are for $|m|$ sufficiently large, normalized in the form

$$(57)\qquad \begin{aligned} u_m(x,\eta) &= \frac{1}{\sqrt{b-a}}\left[e^{2m\pi i\frac{x-a}{b-a}+(x-a)\theta(\eta)}\left\{1+\frac{Q_1(x,\eta)}{m}\right\}\right. \\ &\qquad\left. + \frac{Q_2(x,\eta)}{m} + \frac{H(x,\eta,m)}{m^2}\right], \\ v_m(x,\eta) &= \frac{1}{\sqrt{b-a}}\left[e^{-2m\pi i\frac{x-a}{b-a}-(x-a)\theta(\eta)}\left\{1-\frac{Q_1(x,\eta)}{m}\right\}\right. \\ &\qquad\left. + \frac{Q_3(x,\eta)}{m} + \frac{H(x,\eta,m)}{m^2}\right], \end{aligned}$$

where

$$\theta(\eta) = \frac{1}{b-a}\log\frac{k_1(\eta)}{k_2(\eta)},$$

and the functions $Q_1(x, \eta)$, $Q_2(x, \eta)$, $Q_3(x, \eta)$, $H(x, \eta, m)$ are found from formulas (53), (39), and an analysis of the deductions of Chapters 7, 8 and 9 to be continuous in x and η for η in the region C'''.

It remains to consider the functions of the set $u_m(x, \eta)$, $v_m(x, \eta)$ for small index m. Let us denote by $D(\eta, \lambda)$ the Fredholm determinant for the kernel of equation (56). By the classical theory of integral equations this determinant is analytic in λ and η for η in the region C''', and the characteristic values of equation (56) are the roots $\lambda=\lambda(\eta)$ of the equation

$$(58)\qquad D(\eta,\lambda) = 0.$$

For $\eta=1$ the roots of this equation are all simple by the hypothesis (viii),

and for $\eta=0$ they may be explicitly determined from system (55) and are also found to be simple. It follows* that the equations

$$D(\eta,\lambda) = 0,$$
$$\frac{\partial D(\eta,\lambda)}{\partial\lambda} = 0$$

admit of common roots, which will of course be multiple roots of (58), only for isolated values of η. If we add such of these values as lie in C''' to the set η_i, then for η in any closed sub-region of C''', call it C^{IV}, which excludes also these values, the roots of (58) are all simple and are analytic in η. We observe that by the remarks above $\eta=0$ and $\eta=1$ are not excluded from C^{IV}. We shall designate by $\lambda_m(\eta)$ that root of (58) which joins analytically with $\lambda_m(0)$. It is easily seen that for large values of $|m|$ this agrees with the notation previously adopted in Chapter 8. It follows from the simplicity of the values $\lambda_m(\eta)$ that the entire biorthogonal set $u_m(x, \eta)$, $v_m(x, \eta)$ may be considered normalized. Likewise it is readily found that the functions $u_m(x, \eta)$, $v_m(x, \eta)$ are continuous in x and η.

40. **A change of variable.** The set $u_m(x, \eta)$, $v_m(x, \eta)$ is not directly adaptable to the application of Birkhoff's theorem. For this reason we introduce the set $w_m(x, \eta)$, $z_m(x, \eta)$ defined by the relations

$$w_m(x,\eta) = u_m(x,\eta)e^{-(x-a)\theta(\eta)},$$
$$z_m(x,\eta) = v_m(x,\eta)e^{(x-a)\theta(\eta)}.$$

The points η_i in the circle C form a finite set which includes neither $\eta=0$ nor $\eta=1$. The constant $r>0$ may be so chosen, therefore, that of the circles drawn with radii r and centers η_i no two have a point in common and none includes the point $\eta=0$ or $\eta=1$. We shall suppose such circles drawn and their interiors removed from the circle C. The remaining region is of the type denoted by C^{IV} and in this region we may draw a curve Γ of finite length connecting $\eta=0$ with $\eta=1$. For η on the curve Γ, then, $\theta(\eta)$ is analytic and therefore bounded. Inasmuch as $w_m(x, \eta)\, z_p(x, \eta)=u_m(x, \eta)\, v_p(x, \eta)$, it follows that the set $w_m(x, \eta)$, $z_m(x, \eta)$ is also biorthogonal and normal. From formulas (57) we obtain readily

$$\text{(57a)}\quad \begin{aligned} w_m(x,\eta) &= \frac{1}{\sqrt{b-a}}\left[e^{2m\pi i\frac{x-a}{b-a}}\left\{1+\frac{Q_1(x,\eta)}{m}\right\}+\frac{Q_2'(x,\eta)}{m}+\frac{H(x,\eta,m)}{m^2}\right] \\ z_m(x,\eta) &= \frac{1}{\sqrt{b-a}}\left[e^{-2m\pi i\frac{x-a}{b-a}}\left\{1-\frac{Q_1(x,\eta)}{m}\right\}+\frac{Q_3'(x,\eta)}{m}+\frac{H(x,\eta,m)}{m^2}\right], \end{aligned}$$

* Forsyth, A. R., *Theory of Functions of two Complex Variables*, Cambridge University Press, 1914, pp. 206–209.

in which the functions Q_1, Q_2', Q_3' and H are continuous in x and η, and therefore uniformly continuous for η on Γ and x on the interval (a, b).

41. The application of Birkhoff's theorem. From formulas (57a) we obtain now

$$w_m(x,\eta) - w_m(x,\bar{\eta}) = \frac{1}{b-a}\left[e^{2m\pi i\frac{x-a}{b-a}}\frac{\delta Q_1(x)}{m} + \frac{\delta Q_2'(x)}{m} + \frac{\delta H(x,m)}{m^2}\right]$$

where

$$\delta Q(x) = Q(x,\eta) - Q(x,\bar{\eta}), \quad \text{etc.}$$

It follows then that

$$\begin{aligned}(59)\qquad [w_0(x,\eta) - w_0(x,\bar{\eta})]z_0(y,\eta) + \sum_{m=1}^{\infty}\, &[w_m(x,\eta) - w_m(x,\bar{\eta})]z_m(y,\eta)\\ &+ [w_{-m}(x,\eta) - w_{-m}(x,\bar{\eta})]z_{-m}(y,\eta)\\ = \sum_{m=-M+1}^{M-1}\, &[w_m(x,\eta) - w_m(x,\bar{\eta})]z_m(y,\eta)\\ &+ \frac{\delta Q_1(x)}{b-a}\sum_{m=M}^{\infty}\frac{e^{2m\pi i\frac{x-y}{b-a}} - e^{-2m\pi i\frac{x-y}{b-a}}}{m}\\ &+ \frac{\delta Q_2'(x)}{b-a}\sum_{m=M}^{\infty}\frac{e^{-2m\pi i\frac{y-a}{b-a}} - e^{2m\pi i\frac{y-a}{b-a}}}{m}\\ &+ \frac{1}{b-a}\sum_{m=M}^{\infty}\frac{\delta H(x,m)\cdot H(y,\eta,m)}{m^2}.\end{aligned}$$

Let us consider separately each of the series occurring on the right of this expression. The first series is a finite sum which is small in numerical value for η and $\bar{\eta}$ on Γ and $|\eta-\bar{\eta}|$ small. Also since $H(x, \eta, m)$ is uniformly bounded in m, and is uniformly continuous in x and η for η on Γ and x on (a, b), it follows that the last series converges uniformly in x and y to a value which is numerically small for $|\eta-\bar{\eta}|$ small.

The second and third series may be written respectively in the forms

$$2i\,\frac{\delta Q_1(x)}{b-a}\sum_{m=M}^{\infty}\frac{\sin 2m\pi\left(\dfrac{x-y}{b-a}\right)}{m},$$

$$-\,2i\,\frac{\delta Q_2'(x)}{b-a}\sum_{m=M}^{\infty}\frac{\sin 2m\pi\left(\dfrac{y-a}{b-a}\right)}{m}.$$

Now series of the type $\Sigma[(\sin mz)/m]$ are known to converge uniformly save in the immediate neighborhood of $z=0,\ \pm 2\pi,\ \pm 4\pi, \cdots$, where, however, the sum of any number of terms remains uniformly bounded.* It follows that each of the series in question here converges to a value numerically small because of the uniform smallness of the coefficients $\delta Q_1(x)$, $\delta Q_2'(x)$ for $|\eta-\bar{\eta}|$ small. Moreover when multiplied through by a continuous function of x the series may be integrated term by term with respect to x, and will yield in the case of the former a series which converges uniformly for $a \leqq y \leqq b$, and in the case of the latter a series which converges uniformly on any closed sub-interval of $a<y<b$, and of which the sum of any number of terms remains bounded for $a \leqq y \leqq b$. It is evident, therefore, that for a properly chosen constant $\delta>0$ and $|\eta-\bar{\eta}| \leqq \delta$, η, $\bar{\eta}$ on Γ, the convergence of the series (59) fulfills the hypotheses of Birkhoff's theorem. We may conclude, therefore, that the set of functions $w_m(x, \bar{\eta})$, $z_m(x, \bar{\eta})$ is closed if the set $w_m(x, \eta)$, $z_m(x, \eta)$ is closed.

Let the curve Γ be divided now by points $\eta^{(j)}$ such that

$$\eta^{(0)}=0, \quad |\eta^{(j)}-\eta^{(j-1)}| \leqq \delta,$$

and let the number of divisions be s, so that $\eta^{(s)}=1$. For $\eta^{(0)}$ we have the set of functions

$$w_m(x,0) = \frac{1}{\sqrt{b-a}}\, e^{2m\pi i\frac{x-a}{b-a}}, \qquad z_m(x,0) = \frac{1}{\sqrt{b-a}}\, e^{-2m\pi i\frac{x-a}{b-a}},$$

which is known to be closed. By an application of Birkhoff's theorem we may conclude then that the set $w_m(x, \eta^{(j)})$, $z_m(x, \eta^{(j)})$ is closed for $j=1$, and by successively repeated applications for $j=2, 3, \cdots, s$. This last application proves the closure of the set $w_m(x, 1)$, $z_m(x, 1)$. This, however, implies that the relation

$$\int_a^b u_m(x)f(x)dx = \int_a^b w_m(x,1)e^{\frac{x-a}{b-a}\log\frac{k_1}{k_0}} f(x)dx = 0$$

can be true for all m only if

$$e^{\frac{x-a}{b-a}\log\frac{k_1}{k_0}} f(x) \equiv 0,$$

i.e. if $f(x)\equiv 0$. It follows from this, however, that the set $u_m(x)$, $v_m(x)$ is closed. This is the closure we wished to establish and hence we have the following theorem:

---

* Cf., e. g., Jackson, D., Rendiconti del Circolo Matematico di Palermo, vol. 32 (1911), pp. 257–262, and Bôcher, M., Annals of Mathematics, ser. 2, vol. 7 (1906), pp. 110, 111.

THEOREM 9. *If the kernel of the given integral equation satisfies the conditions* (iv)–(viii), *and that of the associated integral equation satisfies condition* ($\bar{\text{v}}$), *then the set of characteristic solutions* $u_m(x)$, $v_m(x)$ *is closed.*

CHAPTER 11. THE NON-EXISTENCE OF OTHER CHARACTERISTIC VALUES OR FUNCTIONS

42. The possibilities of omission. The deduction of a solution of equation (25a), and hence also the deduction of the characteristic values, was based on the methods of Chapter 6. It is conceivable that other methods might yield a different form of solution, from which either different characteristic values or different characteristic functions or both might result. We shall consider in turn the following questions:

(a) May there exist a characteristic value $\bar{\lambda}$ and a corresponding solution $\bar{u}(x)\not\equiv 0$ which are not included in the sets λ_m and $u_m(x)$?

(b) May there exist a characteristic function $\bar{u}(x)\not\equiv 0$ which corresponds, say, to λ_l, but is not identical with the $u_l(x)$ found?

(c) May the characteristic function $u_l(x)$ found correspond also to a characteristic value $\bar{\lambda}$ not identical with λ_l?

Case (a). Suppose the value $\bar{\lambda}$ and the function $\bar{u}(x)$ mentioned under (a) above to exist. From the biorthogonality of the solutions of associated integral equations it follows that

$$\int_a^b \bar{u}(x)v_m(x)dx = 0 \qquad (m = 0, \pm 1, \pm 2, \cdots).$$

Since the set $u_m(x)$, $v_m(x)$ is closed, however, this demands $\bar{u}(x)\equiv 0$.

Case (b). Suppose $\bar{u}(x)$ is a solution which corresponds to the value λ_l. Then from the biorthogonality we conclude that

$$\int_a^b \bar{u}(x)v_m(x)dx = 0, \qquad m \neq l.$$

If for $m=l$ the integral also vanishes we may conclude as above from the closure of the set $u_m(x)$, $v_m(x)$, that $\bar{u}(x)\equiv 0$. If this is not the case we may suppose the solution $\bar{u}(x)$ normalized by multiplication with a suitable constant so that

$$\int_a^b \bar{u}(x)v_l(x)dx = 1.$$

Then, however,

$$\int_a^b \{\bar{u}(x) - u_l(x)\}v_m(x)dx = 0 \qquad (m = 0, \pm 1, \cdots),$$

and from the closure of the set $u_m(x)$, $v_m(x)$ it follows that

$$\bar{u}(x) \equiv u_l(x).$$

Case (c). Suppose that the function $u_l(x)$ corresponds to a characteristic value $\lambda \neq \lambda_l$. We have in this case

$$u_l(x) = \lambda_l \int_a^b K(x,\xi) u_l(x) dx,$$

$$u_l(x) = \bar{\lambda} \int_a^b K(x,\xi) u_l(x) dx,$$

from which, since $\bar{\lambda} - \lambda_l \neq 0$,

$$\int_a^b K(x,\xi) u_l(\xi) d\xi = 0.$$

This, however, is impossible since $u_l(x)$ is a solution of the given integral equation and $u_l(x) \not\equiv 0$.

It is seen, therefore, that there exist no characteristic values or functions other than those determined in Chapters 8 and 9.

Chapter 12. The expansion of an arbitrary function

43. **The related differential system.** The method by which the properties of the solutions $u_m(x)$ for the expansion of an arbitrary function are to be deduced is based on a suitable comparison of the set $u_m(x)$ with the set of solutions $y_m(x)$ of the differential system

$$\begin{aligned} y'(x) - \rho y(x) &= 0, \\ k_1 y(a) - k_2 y(b) &= 0. \end{aligned} \tag{60}$$

This system is related to the given integral equation in that the constants k_1 and k_2 are those which occur also in the characteristic equation (45).

It is essential, therefore, to have at hand the expansion theorem for a system of type (60). In a form which may be found by specialization of results obtained by the author* for more general differential systems this theorem may be stated as follows:

---

* *Developments associated with a boundary problem not linear in the parameter*, these Transactions, vol. 25 (1923), pp. 155–172. If in the theorem there stated (p. 171) we set $n=1$, $a_1(x)=1$ and change the dependent variable and the parameter respectively by the substitutions

$$u(x) = \sqrt{b-a}\; k_1 y(x)\, e^{\frac{x-a}{b-a}\log\frac{k_2}{k_1}}, \qquad \rho = \frac{1}{\lambda - \alpha_1} - \frac{1}{b-a}\log\frac{k_2}{k_1},$$

the system considered takes the form (60) above.

If $f(x)$ is any function which on the interval (a, b) consists of at most a finite number of pieces, each real and continuous and having a continuous derivative, then $f(x)$ may be expanded in a series of the form

$$f(x) \sim \sum_{m=-\infty}^{+\infty} \int_a^b f(t) z_m(t) dt \cdot y_m(x),$$

where the functions $y_m(x)$ and $z_m(x)$ are respectively the normalized solutions of system (60) above and its adjoint system. This series will converge to $\frac{1}{2}\{f(x+)+f(x-)\}$ for $a<x<b$, to $(1/2k_1)\{k_1 f(a+)+k_2 f(b-)\}$ for $x=a$, and to $(1/2k_2)\{k_1 f(a+)+k_2 f(b-)\}$ for $x=b$.

44. The function $\varphi(x, \xi, p)$. We consider now the function $\varphi(x, \xi, p)$ defined by the relation

$$\varphi(x,\xi,p) = \sum_{m=-p}^{p} [u_m(x)v_m(\xi) - y_m(x)z_m(\xi)]. \tag{61}$$

Its explicit form may be obtained from the solution of system (60), which yields

$$y_m(x)z_m(\xi) = \frac{1}{b-a} e^{\rho_m(x-\xi)},$$

and the formulas (54). It is found in this way that for $|m|$ sufficiently large

$$u_m(x)v_m(\xi) - y_m(x)z_m(\xi) = \frac{1}{b-a}\left\{\frac{Q_1(x) - Q_1(\xi)}{m} e^{\rho_m(x-\xi)} + \frac{Q_2(x)}{m} e^{-\rho_m(\xi-a)} + \frac{H(x,m)H(\xi,m)}{m^2} + \frac{Q_3(x)\, e^{\rho_m(x-a)}}{m}\right\}.$$

With the use, now, of the notation of Chapter 10, i.e.

$$\rho_m = \frac{2m\pi i}{b-a} + \theta, \qquad \theta = \frac{1}{b-a} \log \frac{k_1}{k_2},$$

and with the choice of a number M fixed and sufficiently large, we obtain the formula

$$\varphi(x,\xi,p) = \varphi(x,\xi,M) + \sum_{i=1}^{4} \psi_i(x,\xi,p), \tag{62}$$

where

$$
\begin{aligned}
\psi_1(x,\xi,p) &= \frac{2i}{b-a}\{Q_1(x)-Q_1(\xi)\}e^{\theta(x-\xi)}\sum_{m=M+1}^{p}\frac{\sin 2m\pi\left(\dfrac{x-\xi}{b-a}\right)}{m},\\
\psi_2(x,\xi,p) &= \frac{-2i}{b-a}Q_2(x)e^{-\theta(\xi-a)}\sum_{m=M+1}^{p}\frac{\sin 2m\pi\left(\dfrac{\xi-a}{b-a}\right)}{m},\\
\psi_3(x,\xi,p) &= \sum_{m=M+1}^{p}\frac{H(x,m)H(\xi,m)}{m^2},\\
\psi_4(x,\xi,p) &= \frac{2i}{b-a}Q_3(\xi)e^{\theta(x-a)}\sum_{m=M+1}^{p}\frac{\sin 2m\pi\left(\dfrac{x-a}{b-a}\right)}{m}.
\end{aligned}
\tag{63}
$$

Now sums of the type $\Sigma[(\sin mz)/m]$ as already noted remain uniformly bounded for all values of z. Since this is obviously true also of a sum $\Sigma[H(z,m)/m^2]$ we may conclude that

$$
|\varphi(x,\xi,p)| < A \text{ (a constant)}, \tag{64}
$$

for all x and ξ on (a, b) and all p.

Let α and β be chosen now as any two points of the interval (a, b). We have then from (62) for $p>M$

$$
\int_\alpha^\beta \varphi(x,\xi,p)d\xi = \int_\alpha^\beta \varphi(x,\xi,M)d\xi + \sum_{i=1}^{4}\int_\alpha^\beta \psi_i(x,\xi,p)d\xi. \tag{65}
$$

It is our purpose to consider the convergence of this integral as $p\to\infty$. The first term on the right is free from p and is clearly continuous over the interval (a, b). It is likewise clear that the integral involving ψ_3 will converge uniformly to a function which is continuous on (a, b). For the consideration of the remaining terms we recall the fact, also noted previously, that sums of the type $\Sigma[(\sin mz)/m]$ converge for all values of z and do so uniformly except in the immediate neighborhood of $z=0, \pm 2\pi, \cdots$. It follows because of the integration involved that the integrals containing ψ_1 and ψ_2 converge uniformly over the entire interval, and hence further because of the continuity of the individual terms that the limiting functions are continuous over (a, b).

The integral containing ψ_4 is exceptional. The trigonometric sum involved remains in this case unaffected by the integration, and while this sum converges for all x it does so uniformly only over an interval which does not extend to $x=a$, or $x=b$. Hence the integral converges uniformly

for $a<a_1\leqq x\leqq b_1<b$, and the limiting function is continuous only on the open interval $a<x<b$. We may write, then, upon collecting our results,

(66) $$\lim_{p\to\infty}\int_\alpha^\beta \varphi(x,\xi,p)d\xi = \Phi(x) \qquad \text{uniformly for } a_1 \leqq x \leqq b_1,$$

where $\Phi(x)$ is continuous for $a<x<b$.

45. **The evaluation of** $\Phi(x)$. Let $f(x)$ now be the function defined as follows:

$$f(x) = \begin{cases} 1 \text{ for } \alpha \leqq x \leqq \beta, \\ 0 \text{ for } a \leqq x < \alpha \text{ and } \beta < x \leqq b. \end{cases}$$

This function satisfies the conditions of the theorem in § 43 for expansibility in terms of the solutions $y_m(x)$, whence we have

$$\sum_{m=-\infty}^{+\infty}\int_a^b f(\xi)z_m(\xi)d\xi\cdot y_m(x) = \bar{f}(x),$$

where $\bar{f}(x)$ is bounded and differs from $f(x)$ only in the points α, β, a and b. Inasmuch as

$$\int_a^b f(\xi)\varphi(x,\xi,p)d\xi = \int_\alpha^\beta \varphi(x,\xi,p)d\xi,$$

we obtain with the use of (61) and (66) the result

$$\Phi(x) = \sum_{m=-\infty}^{+\infty}\int_a^b f(\xi)v_m(\xi)d\xi\cdot u_m(x) - \bar{f}(x).$$

Multiplying this equality by $v_l(x)$ and integrating* it with respect to x we find, since the set $u_m(x)$, $v_m(x)$ is normalized biorthogonal, that

$$\int_a^b \Phi(x)v_l(x)dx = \int_a^b f(\xi)v_l(\xi)d\xi - \int_a^b \bar{f}(x)v_l(x)dx = 0.$$

Since this result may be derived for all values of l, whereas the set $u_m(x)$, $v_m(x)$ is closed, it follows that $\Phi(x)\equiv 0$, $a<x<b$. The relation (66) now takes the form

(67) $$\lim_{p\to\infty}\int_\alpha^\beta \varphi(x,\xi,p)d\xi = 0 \qquad \text{uniformly for } a_1 \leqq x \leqq b_1,$$

and this relation holds for all choices of the limits α and β on (a, b).

---

* From the character of the series as shown by the asymptotic forms of $u_m(x)$, $v_m(x)$ this integration term by term is seen to be permissible.

The properties (64) and (67), which have thus been established for the function $\varphi(x, \xi, p)$, are the hypotheses of a theorem by Lebesgue* in accordance with which it follows that if $f(x)$ is any function which is summable over the interval (a, b) then

$$\lim_{p\to\infty}\int_a^b f(\xi)\varphi(x,\xi,p)d\xi = 0 \qquad \text{uniformly for } a_1 \leqq x \leqq b_1. \tag{68}$$

We defer the complete formulation of this result to the end of the chapter.

46. The evaluation of $\Phi(x)$ at the end points of (a, b). It was observed in the derivation of (66) from (65) that the final term of (65) alone fails to converge to a function which is continuous over the entire closed interval (a, b). Accordingly we have

$$\Phi(x)\Big]_{x=b-}^{x=b} = \left[\lim_{p\to\infty}\int_\alpha^\beta \psi_4(x,\xi,p)d\xi\right]_{x=b-}^{x=b}$$

$$= \frac{2i}{b-a}\int_\alpha^\beta Q_3(\xi)d\xi\cdot e^{\theta(b-a)}\left[\sum_{m=M+1}^{\infty}\frac{\sin 2m\pi\left(\dfrac{x-a}{b-a}\right)}{m}\right]_{x=b-}^{x=b}.$$

In virtue of the discussion just completed, however, $\Phi(b-)=0$. On the other hand the series on the right is, except for a finite number of continuous terms, the Fourier expansion of the function $\pi(\frac{1}{2}-(x-a)/(x-b))$. The value of the bracket on the right is, therefore, readily found to be $\pi/2$. Substituting, then, for $Q_3(\xi)$ its value as given by (53b), and for $e^{\theta(b-a)}$ its value k_1/k_2, we obtain the relation

$$\Phi(b) = \frac{1}{2k_2}(\nu k_1 + \mu k_2)\int_\alpha^\beta \bar{\varphi}(\xi)d\xi,$$

which reduces further because of (47) and (44b) to

$$\Phi(b) = \int_\alpha^\beta \Omega_2(\xi)d\xi. \tag{69a}$$

The function $\Omega_2(\xi)$ involved here is that which occurs also in the characteristic equation (45). It is found in an entirely similar fashion that

$$\Phi(a) = \int_\alpha^\beta \Omega_1(\xi)d\xi. \tag{69b}$$

* Annales de la Faculté des Sciences de Toulouse, ser. 3, vol. 1 (1909), p. 52 and p. 68.

Let the functions $\varphi_i(\xi, p)$, $i=1, 2$, be defined now by the relations

$$\varphi_1(\xi,p) = \varphi(a,\xi,p) - \Omega_1(\xi),$$

$$\varphi_2(\xi,p) = \varphi(b,\xi,p) - \Omega_2(\xi).$$

It follows immediately from the bounded character of $\Omega_1(\xi)$ and $\Omega_2(\xi)$, and from the relations (64), (66) and (69) that

$$|\varphi_i(\xi,p)| < A_i \qquad (i = 1,2)$$

for all ξ on (a, b) and all p, and further that

$$\lim_{p\to\infty}\int_\alpha^\beta \varphi_i(\xi,p)d\xi = 0$$

for all choices of α and β on (a, b). By Lebesgue's theorem, then, we may conclude that

$$\lim_{p\to\infty}\int_a^b f(\xi)\varphi_i(\xi,p)d\xi = 0, \tag{70}$$

for every function $f(x)$ which is summable over (a, b).

47. **The expansion theorems.** If we observe now that

$$\int_a^b f(\xi)\varphi(x,\xi,p)d\xi = \sum_{m=-p}^{p}\int_a^b f(\xi)v_m(\xi)d\xi \cdot u_m(x) - \sum_{m=-p}^{p}\int_a^b f(\xi)z_m(\xi)d\xi \cdot y_m(x),$$

we may formulate of the results embodied in the relations (67) and (69) as follows.

THEOREM 10. *If $f(x)$ is any function which is summable over the interval (a, b), and if $F_{1p}(x)$ and $F_{2p}(x)$ are the corresponding sums*

$$F_{1p}(x) = \sum_{m=-p}^{p}\int_a^b f(\xi)v_m(\xi)d\xi \cdot u_m(x),$$

$$F_{2p}(x) = \sum_{m=-p}^{p}\int_a^b f(\xi)z_m(\xi)d\xi \cdot y_m(x),$$

where $u_m(x)$, $v_m(x)$ are respectively the normalized solutions of the given integral equation and its associated equation, and $y_m(x)$, $z_m(x)$ are respectively the normalized solutions of the related differential system (given by (60) *above) and its adjoint system, then*

$$\lim_{p\to\infty} [F_{1p}(x) - F_{2p}(x)] = 0 \qquad \textit{uniformly for } a < a_1 \leqq x \leqq b_1 < b,$$

$$\lim_{p\to\infty} [F_{1p}(a) - F_{2p}(a)] = \int_a^b f(\xi)\Omega_1(\xi)d\xi,$$

$$\lim_{p\to\infty} [F_{1p}(b) - F_{2p}(b)] = \int_a^b f(\xi)\Omega_2(\xi)d\xi.$$

The functions $\Omega_1(\xi)$ *and* $\Omega_2(\xi)$ *are determined by the integral equation and are independent of the function* $f(x)$.

A more explicit but less general formulation becomes possible if we utilize the known expansion theorem for the related differential system as given in § 43. Thus we may state

THEOREM 11. *If* $f(x)$ *is any function which on the interval* (a, b) *consists of at most a finite number of pieces, each real and continuous and having a continuous derivative, then* $f(x)$ *may be expanded in a series of the form*

$$f(x) \sim \sum_{m=-\infty}^{+\infty} \int_a^b f(\xi)v_m(\xi)d\xi \cdot u_m(x).$$

This series converges to

$$\tfrac{1}{2}\{f(x+) + f(x-)\} \qquad \textit{for } a < x < b\,;$$

to

$$\frac{1}{2k_1}\{k_1 f(a+) + k_2 f(b-)\} + \int_a^b f(\xi)\Omega_1(\xi)d\xi \qquad \textit{for } x = a\,;$$

to

$$\frac{1}{2k_2}\{k_1 f(a+) + k_2 f(b-)\} + \int_a^b f(\xi)\Omega_2(\xi)d\xi \qquad \textit{for } x = b.$$

APPENDIX. AN EXAMPLE

It is of interest to apply the preceding theory to a specific example particularly because of the fact that the theory demands in Chapters 8 and 11 that the ratio k_1/k_2 and the functions $\Omega_1(x)$ and $\Omega_2(x)$ be independent of the parameters μ and ν. We consider therefore the integral equation

$$u(x) = \lambda\int_0^b K(x,\xi)u(\xi)d\xi,$$

where

$$K(x,\xi) = \begin{cases} x^2\xi^2 + 2 & \text{for } x \geqq \xi, \\ x^2\xi^2 + 1 & \text{for } x < \xi. \end{cases}$$

The equation is in normal form, and computing the various functions by the respective formulas we find readily

$$\Delta = \mu^2 + (b^4 + 3)\mu\nu + 2\nu^2,$$

$$W(x) = 2\nu + \mu(b^2x^2 + 1),$$

$$V(\xi) = \mu + \nu(b^2\xi^2 + 2),$$

$$E(x,\xi) = \frac{2(\mu + 2\nu)}{\Delta}[(\mu + \nu)x^2 - \nu b^2]\xi,$$

$$D = \frac{(\mu + \nu)[(2 - b^4)\mu + (4 + 2b^4)\nu]}{2\Delta},$$

$$F(x,\xi) = \frac{-E(x,\xi)}{D},$$

$$f(x,\xi) = \frac{-8(\mu + 2\nu)x\xi}{(2 - b^4)\mu + (4 + 2b^4)\nu},$$

$$\varphi(x) = \frac{8b^2x}{(2 - b^4)\mu + (4 + 2b^4)\nu},$$

$$\int_0^b V(\xi)\varphi(\xi)d\xi = \frac{2b^4[2\mu + (4 + b^4)\nu]}{(2 - b^4)\mu + (4 + 2b^4)\nu},$$

$$k_1 = \frac{(2 + b^4)(\mu + \nu)}{D},$$

$$k_2 = \frac{(2 - b^4)(\mu + \nu)}{2D}.$$

The ratio k_1/k_2 is thus free from μ, ν, having the value

$$\frac{k_1}{k_2} = \frac{2(2 + b^4)}{2 - b^4}.$$

Further we find

$$V'(x) + \int_0^b V(t)f(t,x)dt = \frac{-2b^2x(\mu + \nu)}{D},$$

whence

$$\Omega_1(x) = \frac{b^2x}{2 + b^4},$$

$$\Omega_2(x) = \frac{-2b^2x}{2 - b^4}.$$

These values are thus also free from μ and ν as they should be.

DARTMOUTH COLLEGE,
HANOVER, N. H.

THE FIGURATRIX IN THE CALCULUS OF VARIATIONS*

BY

PAUL R. RIDER

Introduction. In the study of a definite integral of the form

$$(1) \qquad \int_{t_1}^{t_2} F[x(t), y(t), x'(t), y'(t)]dt$$

Hadamard defines the *figurative* of the point (x, y) as the curve $F(x', y') = 1$, where x' and y' are the current coördinates, and x and y are considered constant.† The polar reciprocal of the figurative with respect to the unit circle $x'^2 + y'^2 = 1$ is termed by Hadamard the *figuratrix*.‡ The importance of the figuratrix and its space analogue seems to have been overlooked, and it is the desire of the writer to call attention to a number of their interesting properties which enable one to interpret geometrically some of the well known functions and theorems of the calculus of variations.

1. **Figurative, figuratrix, and indicatrix.** We shall prefer to consider an integral of the form

$$(2) \qquad \int_{t_1}^{t_2} f[x(t), y(t), \tau(t)](x'^2(t) + y'^2(t))^{1/2} dt,$$

where $\tau = \arctan(y'/x')$. On account of the homogeneity conditions which must be satisfied by the function F in (1),§ we have

$$(3) \qquad F(x, y, x', y') \equiv F(x, y, \cos\tau, \sin\tau)(x'^2 + y'^2)^{1/2}$$

and the integral (1) can always be reduced to the form (2). In (2), however, the function $f(x, y, \tau)$ need not be periodic in τ; the right member of equation (3) must be. Moreover, the homogeneity conditions mentioned above do not enter into the consideration of the function $f(x, y, \tau)$.||

---

* Presented to the Society, December 26, 1924; received by the editors in December, 1924.

† Hadamard, *Leçons sur le Calcul des Variations*, vol. 1, p. 91.

‡ Hadamard, loc. cit., p. 92.

§ F must be positive homogeneous of degree one in x' and y'. See Bolza, *Vorlesungen über Variationsrechnung*, p. 194.

|| See Bliss, *A new form of the simplest problem of the calculus of variations*, these Transactions, vol. 8 (1907), p. 406. This paper will be referred to as Bliss (I).

Let P, of coördinates x and y, be a fixed point on an extremal* for the integral (2). Take P as an origin, p and q as rectangular coördinates, and consider the envelope of the variable line $p \cos \tau + q \sin \tau = f(\tau)$. It is readily found that the parametric equations of this envelope are

$$(4) \qquad p = f \cos \tau - f_\tau \sin \tau, \quad q = f \sin \tau + f_\tau \cos \tau,$$

functions which play an important rôle in the study of an integral of the form (2). (Subscripts denote partial differentiation.) It is thus easy to identify this envelope with the figuratrix as defined by Hadamard.† For the rectangular coördinates of a point on the figuratrix are $F_{x'}(x', y')$, $F_{y'}(x', y')$. But as a consequence of the homogeneity conditions satisfied by F we have‡

$$(5) \qquad x'F_{x'}(x',y') + y'F_{y'}(x',y') = F(x',y') = F(\cos \tau, \sin \tau)(x'^2 + y'^2)^{1/2},$$

or

$$(6) \qquad F_{x'}(\cos \tau, \sin \tau) \cos \tau + F_{y'}(\cos \tau, \sin \tau) \sin \tau = F(\cos \tau, \sin \tau) = f(\tau).$$

Differentiating, we get

$$(7) \qquad f_\tau(\tau) = -F_{x'}(\cos \tau, \sin \tau) \sin \tau + F_{y'}(\cos \tau, \sin \tau) \cos \tau.$$

Solution of (6) and (7) for $F_{x'}$ and $F_{y'}$ gives

$$(8) \qquad F_{x'} = f \cos \tau - f_\tau \sin \tau, \; F_{y'} = f \sin \tau + f_\tau \cos \tau.$$

The arguments of the functions $F_{x'}$ and $F_{y'}$ may be either (x', y') or $(\cos \tau, \sin \tau)$, since these functions are homogeneous of degree zero with respect to x' and y'. A comparison of (4) and (8) shows that *the envelope of the line* $p \cos \tau + q \sin \tau = f(\tau)$ *is Hadamard's figuratrix.*

If we define the *polar figurative* as the locus

$$(9) \qquad \rho = f(\tau),$$

it is seen that *the polar figurative is the pedal of the figuratrix with respect to the origin P.*

The *indicatrix*§ is the curve

$$(10) \qquad \rho = \frac{1}{f(\tau)} \cdot$$

* See Bolza, loc. cit., pp. 32, 203.

† It may be remarked that the rectangular equation of the figuratrix is of the form $H(x,y,p,q) = 0$, where H is a function satisfying the canonical form of Euler's equations in the calculus of variations. See Hadamard, loc. cit., pp. 151–153.

‡ See Bolza, loc. cit., pp. 194–197.

§ See Bolza, loc. cit., p. 247.

Since $f(\tau)=F(\cos\tau, \sin\tau)=F(x', y')/(x'^2+y'^2)^{1/2}$, equation (10) is equivalent to the equation $F(x', y')=1$. That is, *the indicatrix is the same as the (rectangular) figurative.*

It also follows that

The polar figurative and the indicatrix are inverse curves with respect to the circle* $p^2+q^2=1$;

The figuratrix and the indicatrix are reciprocal polars† with respect to the circle $p^2+q^2=1$.

In Fig. 1, P is the origin (a point on an extremal), Q is a point on the figuratrix, N is a point on the polar figurative, NQ is tangent to the figuratrix.

Since the figuratrix and the indicatrix are reciprocal polars, *the polar of Q* (call it MR) *is tangent to the indicatrix at M.* This tangent line MR, however, gives the direction which is transverse to the direction PN.‡

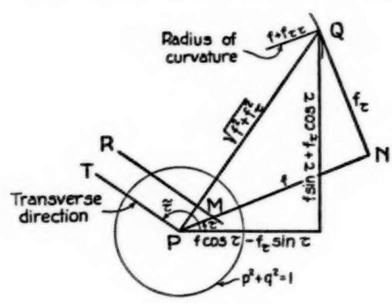

Fig. 1

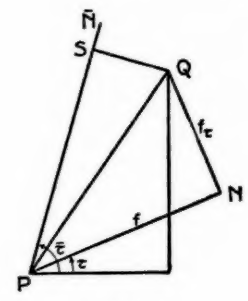

Fig. 2

Thus, *the direction $\bar\tau$ which is transverse to τ, or PN, is the direction PT which is perpendicular to PQ.*

This also follows from the formula§ which gives the direction $\bar\tau$ transverse to τ, viz.:

$$f\cos(\bar\tau-\tau)+f_\tau\sin(\bar\tau-\tau)=0. \tag{11}$$

For a reference to Fig. 1 or Fig. 2 shows that the left member of (11) is simply the projection of PQ upon the line whose direction is $\bar\tau$. When

* See Williamson, *Differential Calculus*, 1899, p. 225.
† See Williamson, loc. cit., p. 228.
‡ See Bolza, loc. cit., p. 304.
§ See Bliss (I), p. 413.

this projection is zero the line having the direction $\bar{\tau}$ is perpendicular to PQ.

*The normal to the polar figurative is the line joining the point N with the middle point of the line PQ.**

2. **Geometric interpretations.** It can be shown without difficulty that $PQ=(f^2+f_{\tau}^2)^{1/2}$. Consequently $NQ=f_\tau$, since $PN=f$. *The radius of curvature of the figuratrix at Q is easily computed to be $f+f_{\tau\tau}$, which is the important f_1-function for an integral of type* (2).

The Weierstrass e-function for the integral (2) is

$$e(x,y,\tau,\bar{\tau}) = f(\bar{\tau}) - [f(\tau)\cos(\bar{\tau}-\tau) + f_\tau(\tau)\sin(\bar{\tau}-\tau)]. \tag{12}$$

If, in Fig. 2, Q is a point on the figuratrix, it is seen immediately that the term in square brackets in (12) is PS, the projection of PQ upon the line making an angle $\bar{\tau}$ with the polar axis. If $P\bar{N}=f(\bar{\tau})$, the e-function is the line segment $S\bar{N}$. If the integral (2) is to be minimized, we must have $e(x, y, \tau, \bar{\tau})\geqq 0$ for the set of values (x, y, τ) giving the point P and the direction of the extremal at that point, and for all values of $\bar{\tau}$.† Consequently, *the projection of PQ upon the line $\tau=\bar{\tau}$ cannot be greater than $f(\bar{\tau})$.*

If d is the distance from the tangent line to the figuratrix at $Q(\tau)$, i. e. $p\cos\tau+q\sin\tau-f(\tau)=0$, to the point $\bar{Q}(\bar{\tau})$ on the curve, then

$$\begin{aligned} d &= [f(\bar{\tau})\cos\bar{\tau} - f_\tau(\bar{\tau})\sin\bar{\tau}]\cos\tau + [f(^-)\sin\bar{\tau} + f_\tau(\bar{\tau})\cos\bar{\tau}]\sin\tau - f(\tau) \\ &= f(\bar{\tau})\cos(\bar{\tau}-\tau) - f_\tau(\bar{\tau})\sin(\bar{\tau}-\tau) - f(\tau). \end{aligned}$$

Expanding in powers of $(\bar{\tau}-\tau)$ we find that

$$\begin{aligned} d &= -\tfrac{1}{2}(\bar{\tau}-\tau)^2[f(\tau)+f_{\tau\tau}(\tau)] + h \\ &= -\tfrac{1}{2}(\bar{\tau}-\tau)^2 f_1 + h. \end{aligned} \tag{13}$$

A similar expansion gives

$$\begin{aligned} e(\tau,\bar{\tau}) &= f(\bar{\tau}) - [f(\tau)\cos(\bar{\tau}-\tau) + f_\tau(\tau)\sin(\bar{\tau}-\tau)] \\ &= \tfrac{1}{2}(\bar{\tau}-\tau)^2[f(\tau)+f_{\tau\tau}(\tau)] + j \\ &= \tfrac{1}{2}(\bar{\tau}-\tau)^2 f_1 + j. \end{aligned}$$

Thus

$$e(\tau,\bar{\tau}) = -d + k. \tag{14}$$

---

* See Goursat-Hedrick, *Mathematical Analysis*, vol. 1, p. 69.

† See Bliss (I), p. 412.

The quantities h, j, k are homogeneous of degree three in $(\bar{\tau}-\tau)$ and contain partial derivatives of f with respect to τ to an order not exceeding the fourth. If these derivatives are finite, d and e are infinitesimals of the second order with respect to $(\bar{\tau}-\tau)$. The geometric significance of e is apparent from (14). *The function* $e(\tau, \bar{\tau})$ *is the distance from the point* $\bar{Q}(\bar{\tau})$ *on the figuratrix to the tangent line at the point* $Q(\tau)$, *except for an infinitesimal of at least the third order with respect to* $(\bar{\tau}-\tau)$. Under the assumption $f_1 \geqq 0$,* this distance is positive† for points $\bar{Q}$ near Q, and consequently *the figuratrix is concave toward the origin P at the point* $Q(\tau)$.

A well known theorem due to Hedrick states that all integrand functions for which the transversality condition is equivalent to orthogonality are of the form $F(x, y, x', y') = G(x, y)\,(x'^2+y'^2)^{1/2}$. In § 1 we found that the direction transverse to $PN(\tau)$ in Fig. 1 is $PT(\bar{\tau})$, which is perpendicular to PQ. But if transversality is to be orthogonality, that is, if PT is to be perpendicular to PN, then PN and PQ must coincide. This of course means that the figuratrix, which it will be remembered is the envelope of the variable line NQ, must coincide with the polar figurative (locus of N) and that both curves must be circles. Therefore the equation of the polar figurative must be independent of τ, which means that the function $f(x, y, \tau)$ must be of the form $g(x, y)$. Thus we have a simple proof of Hedrick's theorem.

We shall now investigate the case of discontinuous solutions.‡ Let P_1PP_2 be an extremal having a corner at the point P. Let τ be the direction of P_1P at P, and $\bar{\tau}$ the direction of PP_2 at P. Then at P the following relations, called *corner conditions*, must be satisfied:§

$$f(\tau)\cos\tau - f_\tau(\tau)\sin\tau = f(\bar{\tau})\cos\bar{\tau} - f_\tau(\bar{\tau})\sin\bar{\tau},$$
$$f(\tau)\sin\tau + f_\tau(\tau)\cos\tau = f(\bar{\tau})\sin\bar{\tau} + f_\tau(\bar{\tau})\cos\bar{\tau}.$$

If we refer to (4), the parametric equations for the figuratrix, we see that the above corner conditions simply state that *for the corner P the directions τ and $\bar{\tau}$ give the same point Q on the figuratrix.* This is to be expected, since corresponding to a corner point on the extremal we have a double tangent to the indicatrix,‖ and the figuratrix is the reciprocal polar of the indicatrix.

* The function f_1 must be $\geqq 0$ if the integral in (2) is to be minimized. See Bliss (I), p. 409.
† Since it is the negative of d.
‡ See Bolza, loc. cit., chapter 8.
§ See Rider, *A note on discontinuous solutions in the calculus of variations*, Bulletin of the American Mathematical Society, vol. 23 (1917), p. 239.
‖ See Bolza, loc. cit., p. 369.

It follows that *the directions which are transverse to τ and $\bar{\tau}$ at the corner P coincide*,* since the transverse direction is obtained by drawing a perpendicular to PQ.

Bliss has defined generalized angle† and generalized area‡ as

$$(15) \qquad \int \frac{f(\tilde{\tau})}{f^2(\tau)}\,(f^2(\tau)+f_\tau^2(\tau))^{1/2}d\tau, \qquad \int\int f(\tilde{\tau})(f^2(\tau)+f_\tau^2(\tau))^{1/2}dx\,dy,$$

respectively, in which $\tilde{\tau}$ is the direction transverse to τ. It may be worth mentioning that if in Fig. 1 we mark off on PT a length $P\tilde{N}$ equal to $f(\tilde{\tau})$, then *the integrand of the second integral in* (15) *is the area of the rectangle whose* sides are PQ and $P\tilde{N}$, and the integrand of the first integral is the ratio of the area of this rectangle to the area of the square whose side is PN.

3. **The figuratrix in space.** For an integral of the form

$$(16) \qquad \int_{t_1}^{t_2} F(x,y,z,x',y',z')dt,$$

in which x, y, z are functions of t, the figurative of the point (x, y, z) is defined§ as the surface $F(x', y', z')=1$. The figuratrix is defined as the polar reciprocal of this surface with respect to the unit sphere $x'^2+y'^2+z'^2=1$. We shall, however, deal with the integral||

$$(17) \qquad \int_{t_1}^{t_2} f(x,y,z,\tau,\sigma)(x'^2+y'^2+z'^2)^{1/2}dt,$$

where τ and σ are defined by the equations

$$(18) \qquad \begin{aligned} x'/(x'^2+y'^2+z'^2)^{1/2} &= \cos\tau\cos\sigma,\\ y'/(x'^2+y'^2+z'^2)^{1/2} &= \sin\tau\cos\sigma,\\ z'/(x'^2+y'^2+z'^2)^{1/2} &= \sin\sigma. \end{aligned}$$

The geometric significance of τ and σ is at once evident. If at any point of a space curve the positive tangent is constructed, then τ is the angle which this tangent line makes with its own projection in the xy-plane, and σ is the angle which this projection makes with the x-axis. The figuratrix will be defined as the envelope of the two-parameter family of planes

$$(19) \qquad p\cos\tau\cos\sigma + q\sin\tau\cos\sigma + r\sin\sigma = f(\tau,\sigma),$$

* Cf. Bolza, loc. cit., p. 369.

† Bliss, *A generalization of the notion of angle*, these Transactions, vol. 7 (1906), pp. 184–196.

‡ Bliss, *Generalizations of geodesic curvature and a theorem of Gauss concerning geodesic triangles*, American Journal of Mathematics, vol. 37 (1915), pp. 1–18.

§ See Hadamard, loc. cit., p. 96.

|| Cf. Rider, *The space problem of the calculus of variations in terms of angle*, American Journal of Mathematics, vol. 39 (1917), pp. 241–256. This paper will be referred to as Rider (II).

p, q, r being rectangular coördinates with respect to an origin P on an extremal.

As before, this envelope can be proved identical with Hadamard's figuratrix. In this section we shall develop some of its properties, which are perhaps even more interesting than those of the figuratrix in two dimensions. Where the generalization from two to three dimensions is quite obvious, proofs will be omitted.

By differentiating equation (19) partially with respect to τ and σ we obtain the equations

$$(20)\qquad \begin{aligned} -p\sin\tau\cos\sigma + q\cos\tau\cos\sigma &= f_\tau(\tau,\sigma),\\ -p\cos\tau\sin\sigma - q\sin\tau\sin\sigma + r\cos\sigma &= f_\sigma(\tau,\sigma). \end{aligned}$$

Solving equations (19) and (20) for p, q, r (this can be done if $\cos\sigma\neq 0$), we get*

$$(21)\qquad \begin{aligned} p &= f\cos\tau\cos\sigma - f_\tau\sin\tau/\cos\sigma - f_\sigma\cos\tau\sin\sigma,\\ q &= f\sin\tau\cos\sigma + f_\tau\cos\tau/\cos\sigma - f_\sigma\sin\tau\sin\sigma,\\ r &= f\sin\sigma + f_\sigma\cos\sigma, \end{aligned}$$

the rectangular coördinates of a point Q on the figuratrix. It is readily verified that $PQ=(p^2+q^2+r^2)^{1/2}=(f^2+f_\tau^2/\cos^2\sigma+f_\sigma^2)^{1/2}$.

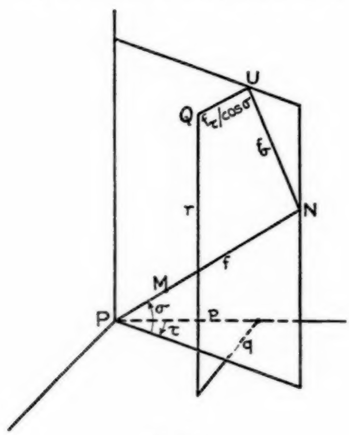

Fig. 3

Let $PN=f(\tau,\sigma)$, $PM=1/f(\tau,\sigma)$. (See Fig. 3.) Draw $NU=f_\sigma$ making an angle of $\pi/2$ with PN and lying in the plane of PN and the r-axis. Then

* Cf. Rider (II), p. 243, equations (5).

the line-segment UQ will be parallel to the pq-plane, perpendicular to NU, and equal in length to $f_\tau/\cos\sigma$. Moreover

The polar figurative is the pedal surface of the figuratrix with respect to the origin P;

The indicatrix is identical with the (rectangular) figurative;

The polar figurative and the indicatrix are inverse surfaces with respect to the sphere $p^2+q^2+r^2=1$;

The figuratrix and the indicatrix are reciprocal polars with respect to the sphere $p^2+q^2+r^2=1$;

The plane NUQ is of course tangent at Q to the figuratrix; the plane through M perpendicular to PQ is tangent to the indicatrix at M;

The normal to the polar figurative is the line joining the point N with the middle point of the line PQ.

The direction $PT(\tilde\tau, \tilde\sigma)$ which is transverse to $PN(\tau, \sigma)$ is given by the equation*

$$\begin{aligned} f(\cos\tau\cos\sigma\cos\tilde\tau\cos\tilde\sigma + \sin\tau\cos\sigma\sin\tilde\tau\cos\tilde\sigma + \sin\sigma\sin\tilde\sigma) \\ + f_\tau\left(\frac{\cos\tau}{\cos\sigma}\sin\tilde\tau\cos\tilde\sigma - \frac{\sin\tau}{\cos\sigma}\cos\tilde\tau\cos\tilde\sigma\right) \\ - f_\sigma(\cos\tau\sin\sigma\cos\tilde\tau\cos\tilde\sigma + \sin\tau\sin\sigma\sin\tilde\tau\cos\tilde\sigma - \cos\sigma\sin\tilde\sigma) = 0. \end{aligned}$$

Since the direction cosines of $PT(\tilde\tau, \tilde\sigma)$ are $\cos\tilde\tau\cos\tilde\sigma$, $\sin\tilde\tau\cos\tilde\sigma$, $\sin\tilde\sigma$, and those of PQ are proportional to p, q, r of (21), it is seen that *the direction $PT(\tilde\tau, \tilde\sigma)$ which is transverse to the direction $PN(\tau, \sigma)$ is perpendicular to PQ.*

It follows at once that *if transversality means orthogonality, the function $f(x, y, z, \tau, \sigma)$ must be of the form $g(x, y, z)$.*

The parameters of the surface (21) are τ and σ and the fundamental quantities of the first order are

$$E = p_\tau^2 + q_\tau^2 + r_\tau^2, \qquad F = p_\tau p_\sigma + q_\tau q_\sigma + r_\tau r_\sigma, \qquad G = p_\sigma^2 + q_\sigma^2 + r_\sigma^2.$$

The functions p_τ, q_τ, r_τ, p_σ, q_σ, r_σ are the quantities $-a_1$, $-a_2$, $-a_3$ $-b_1$, $-b_2$, $-b_3$, respectively, found in an earlier paper by the writer,† and it is not difficult to show that

---

* Cf. Rider (II), p. 250.

† Rider (II), pp. 243–244.

$$E = (f_\tau \tan\sigma + f_{\tau\sigma})^2/\cos^2\sigma + (f + f_{\sigma\sigma})^2,$$
$$F = (f_\tau \tan\sigma + f_{\tau\sigma})[(f + f_{\sigma\sigma}) + (f - f_\sigma \tan\sigma + f_{\tau\tau}/\cos^2\sigma)],$$
$$G = (f_\tau \tan\sigma + f_{\tau\sigma})^2 + \cos^2\sigma(f - f_\sigma \tan\sigma + f_{\tau\tau}/\cos^2\sigma)^2,$$
$$H = (EG - F^2)^{1/2} = \pm f_1/\cos\sigma,$$

where

$$f_1 = (f\cos^2\sigma - f_\sigma \sin\sigma\cos\sigma + f_{\tau\tau})(f + f_{\sigma\sigma}) - (f_\tau \tan\sigma + f_{\tau\sigma})^2.$$

If

$$X = \frac{1}{H}\begin{vmatrix} q_\tau & r_\tau \\ q_\sigma & r_\sigma \end{vmatrix}, \quad Y = \frac{1}{H}\begin{vmatrix} r_\tau & p_\tau \\ r_\sigma & p_\sigma \end{vmatrix}, \quad Z = \frac{1}{H}\begin{vmatrix} p_\tau & q_\tau \\ p_\sigma & q_\sigma \end{vmatrix},$$

then $X = \cos\tau\cos\sigma$, $Y = \sin\tau\cos\sigma$, $Z = \sin\sigma$. This result is to be expected; for X, Y, Z are the direction cosines of the normal to the figuratrix, and the plane (19) is tangent to the figuratrix. The fundamental coefficients of the second order for the figuratrix are

$$\begin{aligned} D &= Xp_{\tau\tau} + Yq_{\tau\tau} + Zr_{\tau\tau} = -(X_\tau p_\tau + Y_\tau q_\tau + Z_\tau r_\tau) \\ &= -(f\cos^2\sigma - f_\sigma \sin\sigma\cos\sigma + f_{\tau\tau}), \\ D' &= Xp_{\tau\sigma} + Yq_{\tau\sigma} + Zr_{\tau\sigma} = -(X_\tau p_\sigma + Y_\tau q_\sigma + Z_\tau r_\sigma) \\ &= -(X_\sigma p_\tau + Y_\sigma q_\tau + Z_\sigma r_\tau) = -(f_\tau \tan\sigma + f_{\tau\sigma}), \\ D'' &= Xp_{\sigma\sigma} + Yq_{\sigma\sigma} + Zr_{\sigma\sigma} = -(X_\sigma p_\sigma + Y_\sigma q_\sigma + Z_\sigma r_\sigma) \\ &= -(f + f_{\sigma\sigma}). \end{aligned}$$

The total curvature of the figuratrix is $(DD'' - D'^2)/H^2 = \cos^2\sigma/f_1$. Since for a regular problem f_1 must be positive,* *the figuratrix is a surface of positive curvature at* Q.

It can be shown that if the integral (17) is to be minimized or maximized the function f_1 cannot be negative. If we assume $f_1 > 0$, then D and D'' must be negative for a minimum and positive for a maximum.†

The Weierstrass e-function for the integral (17) is‡

$$\begin{aligned} e(\tau,\sigma,\bar\tau,\bar\sigma,) &= f(\bar\tau,\bar\sigma) \\ &\quad - [f(\tau,\sigma,)(\cos\tau\cos\sigma\cos\bar\tau\cos\bar\sigma + \sin\tau\cos\sigma\sin\bar\tau\cos\bar\sigma + \sin\sigma\sin\bar\sigma) \\ &\quad + \frac{f_\tau(\tau,\sigma)}{\cos\sigma}\cos\bar\sigma(-\sin\tau\cos\bar\tau + \cos\tau\sin\bar\tau) \\ &\quad + f_\sigma(\tau,\sigma)(-\cos\tau\sin\sigma\cos\bar\tau\cos\bar\sigma - \sin\tau\sin\sigma\sin\bar\tau\cos\bar\sigma \\ &\quad + \cos\sigma\sin\bar\sigma)]. \end{aligned} \tag{22}$$

* See Rider (II), p. 249.

† See Rider (II), p. 249.

‡ See Rider (II), p. 247.

But PS, the projection of the line PQ upon the line whose direction is $(\bar{\tau}, \bar{\sigma})$, is equal to

$$p \cos \bar{\tau} \cos \bar{\sigma} + q \sin \bar{\tau} \cos \bar{\sigma} + r \sin \bar{\sigma}, \tag{23}$$

in which p, q, r are given by (21). It is not difficult to show that the expression (23) is the same as that in the square brackets in (22). Therefore, if $P\bar{N} = f(\bar{\tau}, \bar{\sigma})$ is marked off on the line whose direction is given by $(\bar{\tau}, \bar{\sigma})$, the e-function is the line segment $S\bar{N}$. If the integral (17) is to be made a minimum, we must have $e(x, y, z, \tau, \sigma, \bar{\tau}, \bar{\sigma}) \geqq 0$ for the set of values (x, y, z, τ, σ) giving the point P and the direction of the extremal at that point, and for all values of $\bar{\tau}$ and $\bar{\sigma}$.* Consequently, *the projection of PQ upon the line* $\tau = \bar{\tau}$, $\sigma = \bar{\sigma}$ *cannot be greater than* $f(\bar{\tau}, \bar{\sigma})$.

The e-function may be written in the form†

$$e(\tau, \sigma, \bar{\tau}, \bar{\sigma}) = -\tfrac{1}{2}\left[(\bar{\tau} - \tau)^2 D + 2(\bar{\tau} - \tau)(\bar{\sigma} - \sigma) D' + (\bar{\sigma} - \sigma)^2 D''\right] + h.$$

Thus,‡ *the e-function is the distance from the point* $\bar{Q}(\bar{\tau}, \bar{\sigma})$ *on the figuratrix to the tangent plane at the point* $Q(\tau, \sigma)$, *except for an infinitesimal of at least the third order with respect to* $(\bar{\tau} - \tau)$ *and* $(\bar{\sigma} - \sigma)$.

Since the e-function must be positive for a minimum of the integral (17), *the figuratrix is concave at Q toward the origin P.*

The statements regarding discontinuous solutions admit of obvious generalization from two to three dimensions.

4. **Generalization to n dimensions.** For the sake of completeness we shall generalize for the case of n dimensions. Let us consider the integral§

$$\int_{t_1}^{t_2} f(x_1, \cdots, x_n, \tau_1, \cdots, \tau_{n-1})(x_1'^2 + \cdots + x_n'^2)^{1/2} dt, \tag{24}$$

in which the x's and τ's are functions of t. The quantity τ_{n-1} is the angle made by the positive tangent to a curve with its own projection in the $(n-1)$-space $(x_1, \cdots, x_{n-1})$, τ_{n-2} is the angle made by this projection with its own projection in the $(n-2)$-space $(x_1, \cdots, x_{n-2})$, and so on. Analytically the τ's are defined by the equations

$$\tau_i = \arctan\left(x_{i+1}'/(x_1'^2 + \cdots + x_i'^2)^{1/2}\right) \qquad (i = 1, \cdots, n-1), \tag{25}$$

* See Rider (II), p. 248.

† See Rider (II), p. 248.

‡ See Eisenhart, *Differential Geometry*, p. 114.

§ Cf. Rider, *On the problem of the calculus of variations in n dimensions*, Tôhoku Mathematical Journal, vol. 13 (1918), pp. 165–171. This paper will be referred to as Rider (III).

from which it is seen that

$$x_i'/(x_1'^2 + \cdots + x_n'^2)^{1/2} = \sin\tau_{i-1}\cos\tau_i\cdots\cos\tau_{n-1}\ (i = 1, \cdots, n-1) \tag{26}$$

(if we make the convention $\sin\tau_0 = 1$).

Let us now take as origin a point $P(x_1, \cdots, x_n)$ on an extremal for the integral (24), and as coördinates the variables $p_1, \cdots, p_n$. We shall define the figuratrix of the point P as the envelope of the variable hyperplane

$$\sum_{i=1}^{n} p_i \sin\tau_{i-1}\cos\tau_i\cdots\cos\tau_{n-1} = f(\tau_1, \cdots, \tau_{n-1}). \tag{27}$$

Differentiating partially with respect to $\tau_1, \cdots, \tau_{n-1}$, we obtain the following equations:

$$\begin{aligned} &-p_1\sin\tau_1\cos\tau_2\cdots\cos\tau_{n-1} + p_2\cos\tau_1\cdots\cos\tau_{n-1} = f_{\tau_1},\\ &-p_1\cos\tau_1\sin\tau_2\cos\tau_3\cdots\cos\tau_{n-1} - p_2\sin\tau_1\sin\tau_2\cos\tau_3\cdots\cos\tau_{n-1}\\ &\qquad + p_3\cos\tau_3\cdots\cos\tau_{n-1} = f_{\tau_2},\\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ &-p_1\cos\tau_1\cdots\cos\tau_{n-2}\sin\tau_{n-1} - \cdots - p_{n-1}\sin\tau_{n-2}\sin\tau_{n-1}\\ &\qquad + p_n\cos\tau_{n-1} = f_{\tau_{n-1}}. \end{aligned} \tag{28}$$

If $\cos\tau_2\cos^2\tau_3\cdots\cos^{n-2}\tau_{n-1}$, the determinant of the coefficients of the system of equations composed of (27) and (28), is different from zero we can solve for the p's obtaining

$$\begin{aligned} p_i = f\sin\tau_{i-1}\cos\tau_i\cdots\cos\tau_{n-1} + f_{\tau_{i-1}}\frac{\cos\tau_{i-1}}{\cos\tau_i\cdots\cos\tau_{n-1}}&\\ - \sum_{j=i}^{n-1} f_{\tau_j}\frac{\sin\tau_j\sin\tau_{i-1}\cos\tau_i\cdots\cos\tau_{j-1}}{\cos\tau_{j+1}\cdots\cos\tau_{n-1}}&\qquad (i = 1, \cdots, n), \end{aligned} \tag{29}$$

the rectangular coördinates of a point Q on the figuratrix.

As in the preceding cases, *the direction which is transverse to the direction* $(\tau_1, \cdots, \tau_{n-1})$ *is orthogonal to* PQ.*

Since the normal to the figuratrix is parallel to the line PN, the direction cosines of this normal are

$$X_i = x_i'/(x_1'^2 + \cdots + x_n'^2)^{1/2} = \sin\tau_{i-1}\cos\tau_i\cdots\cos\tau_{n-1}\ (i = 1, \cdots, n). \tag{30}$$

---

* For the transversality condition see Rider (III), pp. 168–169.

The fundamental coefficients of the second order for the figuratrix are then defined as

$$D_{jk} = - \sum_{i=1}^{n} \frac{\partial X_i}{\partial \tau_j} \frac{\partial p_i}{\partial \tau_k}. \tag{31}$$

The partial derivatives occurring in D_{jk} are found to be

$$\frac{\partial X_i}{\partial \tau_j} = - \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{j-1} \sin \tau_j \cos \tau_{j+1} \cdots \cos \tau_{n-1} \qquad (j \geqq i),$$

$$\frac{\partial X_i}{\partial \tau_{i-1}} = \cos \tau_{i-1} \cos \tau_i \cdots \cos \tau_{n-1},$$

$$\frac{\partial X_i}{\partial \tau_j} = 0 \qquad (j < i-1),$$

$$\begin{aligned}
\frac{dp_i}{d\tau_j} = &- \frac{\sin \tau_j \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{j-1}}{\cos \tau_{j+1} \cdots \cos \tau_{n-1}} \Big(f \cos^2 \tau_{j+1} \cdots \cos^2 \tau_{n-1} \\
&- \sum_{k=j+1}^{n-1} f_{\tau_k} \sin \tau_k \cos \tau_k \cos^2 \tau_{j+1} \cdots \cos^2 \tau_{k-1} + f_{\tau_j \tau_j} \Big) \\
&+ \frac{\cos \tau_{i-1}}{\cos \tau_i \cdots \cos \tau_{n-1}} \left(f_{\tau_{i-1}} \tan \tau_j + f_{\tau_{i-1} \tau_j} \right) \\
&- \sum_{k=i}^{j-1} \frac{\sin \tau_k \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{k-1}}{\cos \tau_{k+1} \cdots \cos \tau_{n-1}} \left(f_{\tau_k} \tan \tau_j + f_{\tau_k \tau_j} \right) \\
&- \sum_{k=j+1}^{n-1} \frac{\sin \tau_k \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{k-1}}{\cos \tau_{k+1} \cdots \cos \tau_{n-1}} \left(f_{\tau_j} \tan \tau_k + f_{\tau_j \tau_k} \right) \quad (j \geqq i),
\end{aligned}$$

$$\begin{aligned}
\frac{dp}{d\tau_{i-1}} = &\frac{\cos \tau_{i-1}}{\cos \tau_i \cdots \cos \tau_{n-1}} \Big(f \cos^2 \tau_i \cdots \cos^2 \tau_{n-1} \\
&- \sum_{k=i}^{n-1} f_{\tau_k} \sin \tau_k \cos \tau_k \cos^2 \tau_i \cdots \cos^2 \tau_{k-1} + f_{\tau_{i-1} \tau_{i-1}} \Big) \\
&- \sum_{k=i}^{n-1} \frac{\sin \tau_k \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{k-1}}{\cos \tau_{k+1} \cdots \cos \tau_{n-1}} \left(f_{\tau_{i-1}} \tan \tau_k + f_{\tau_{i-1} \tau_k} \right),
\end{aligned}$$

$$\begin{aligned}
\frac{dp_i}{d\tau_j} = &\frac{\cos \tau_{i-1}}{\cos \tau_i \cdots \cos \tau_{n-1}} \left(f_{\tau_j} \tan \tau_{i-1} + f_{\tau_j \tau_{i-1}} \right) \\
&- \sum_{k=i}^{n-1} \frac{\sin \tau_k \sin \tau_{i-1} \cos \tau_i \cdots \cos \tau_{k-1}}{\cos \tau_{k+1} \cdots \cos \tau_{n-1}} \left(f_{\tau_j} \tan \tau_k + f_{\tau_j \tau_k} \right) \quad (j < i-1).
\end{aligned}$$

It is found that

$$D_{ii} = -\Big(f\cos^2\tau_{i+1}\cdots\cos^2\tau_{n-1} - \sum_{j=i+1}^{n-1} f_{\tau_j}\sin\tau_j\cos\tau_j\cos^2\tau_{i+1}\cdots\cos^2\tau_{j-1}$$

$$+ f_{\tau_i\tau_i}\Big),$$

$$D_{ij} = -\left(f_{\tau_i}\tan\tau_j + f_{\tau_i\tau_j}\right) \qquad (i \neq j).$$

(It is assumed that $f\cos^2\tau_{i+1}\cdots\cos^2\tau_{n-1}=f$ when $i=n-1$.) The writer has already shown that *for a minimum of the integral* (24) *the quantities D_{ii} cannot be positive.**

Let us define the quantity H by means of the formula

$$H^2 = \left\| \begin{matrix} \dfrac{\partial p_1}{\partial \tau_1} \cdots \dfrac{\partial p_1}{\partial \tau_{n-1}} \\ \cdots\cdots\cdots \\ \dfrac{\partial p_n}{\partial \tau_1} \cdots \dfrac{\partial p_n}{\partial \tau_{n-1}} \end{matrix} \right\|^2 = \sum_{i=1}^{n} D_i^2,$$

where D_i is the determinant obtained by omitting the ith row from the matrix. Then†

$$D_i = (-1)^{i-1} f_1 X_i / \cos\tau_2\cos^2\tau_3\cdots\cos^{n-2}\tau_{n-1},$$

where f_1 is the symmetric determinant

$$f_1 = (-1)^{n-1}\left| \begin{matrix} D_{11} \cdots D_{1n} \\ \cdots\cdots\cdots \\ D_{1n} \cdots D_{nn} \end{matrix} \right|.$$

Thus

$$H^2 = f_1^2/(\cos\tau_2\cos^2\tau_3\cdots\cos^{n-2}\tau_{n-1})^2.$$

The total curvature of the figuratrix is

$$\frac{1}{H^2}\left| \begin{matrix} D_{11} \cdots D_{1n} \\ \cdots\cdots\cdots \\ D_{1n} \cdots D_{nn} \end{matrix} \right| = \frac{(-1)^{n-1}}{f_1}(\cos\tau_2\cos^2\tau_3\cdots\cos^{n-2}\tau_{n-1})^2.$$

---

* See Rider (III), p. 171. In that paper α_{ij} is the negative of D_{ij}.

† See Rider, *On the f_1-function in the calculus of variations*, Washington University Studies, vol. 5, scientific series no. 2 (Jan., 1918), p. 100.

The Weierstrass e-function is the function*

$$e(\tau_1, \cdots, \tau_{n-1}, \bar{\tau}_1, \cdots, \bar{\tau}_{n-1}) = f(\bar{\tau}_1, \cdots, \bar{\tau}_{n-1}) - \sum_{i=1}^{n} p_i \sin \bar{\tau}_{i-1} \cos \bar{\tau}_i \cdots \cos \bar{\tau}_{n-1}$$

$$= -\frac{1}{2}\left[\sum_{i=1}^{n-1} (\bar{\tau}_i - \tau_i)^2 D_{ii} + \sum_{i=1,\, j=i+1}^{n-1} (\bar{\tau}_i - \tau_i)(\bar{\tau}_j - \tau_j) D_{ij} \right] + h,$$

the p's being defined by (29). It thus appears, that with proper restrictions on the partial derivatives of f which occur in h, *the e-function is the distance from the point $\bar{Q}$ on the figuratrix to the tangent hyperplane at Q except for an infinitesimal of at least the third order with respect to the quantities* $(\bar{\tau}_i - \tau_i)$. The concavity of the figuratrix toward the origin P follows, since e must be positive for a minimum value of the integral (24).

* See Rider (III), pp. 169–170.

WASHINGTON UNIVERSITY
ST. LOUIS, MO.

ON THE EXISTENCE OF FIELDS IN BOOLEAN ALGEBRAS*

BY

B. A. BERNSTEIN

A class K is said to be a *field* with respect to a pair of operations Δ, $\bigcirc$ if Δ, $\bigcirc$ play in K the same rôle which the operations $+$, $\times$ play in the class of rational numbers. My aim in this paper is to determine in any boolean algebra all pairs of operations expressible in terms of addition, multiplication, and negation for which the elements are a field.

Postulates for fields. The conditions that make a class K a field with respect to a pair of operations Δ, $\bigcirc$ are† given by the following seven postulates:‡

P_1. $x\Delta y = y\Delta x$, if x, y, $y\Delta x$ are K-elements.

P_2. $(x\Delta y)\Delta z = x\Delta(y\Delta z)$, if x, y, z, $x\Delta y$, $y\Delta z$, $x\Delta(y\Delta z)$ are K-elements.

P_3. For any two K-elements x, y there is a K-element ξ such that $x\Delta\xi = y$.

P_4. $x\bigcirc y = y\bigcirc x$, if x, y, $y\bigcirc x$ are K-elements.

P_5. $(x\bigcirc y)\bigcirc z = x\bigcirc(y\bigcirc z)$, if x, y, z, $x\bigcirc y$, $y\bigcirc z$, $x\bigcirc(y\bigcirc z)$ are K-elements.

P_6. $x\bigcirc(y\Delta z) = (x\bigcirc y)\Delta(x\bigcirc z)$, if x, y, z, $y\Delta z$, $x\bigcirc y$, $x\bigcirc z$, $(x\bigcirc y)\Delta(x\bigcirc z)$ are K-elements.

P_7. For any two K-elements x, y satisfying the conditions $x\Delta x \neq x$, $y\Delta y \neq y$, there is a K-element η such that $x\bigcirc\eta = y$.

Conditions on boolean operations imposed by the postulates. In a boolean algebra any binary operation expressible in terms of addition, multiplication, and negation must be of the form

$$Axy + Bxy' + Cx'y + Dx'y',\S$$

where A, B, C, D are elements of the algebra. Let two boolean operations Δ, $\bigcirc$ be given by

---

* Presented to the Society, San Francisco Section, April 4, 1925; received by the editors in December, 1925.

† Of a pair of field operations, the operation appearing first will be the one possessing the properties that "+" has in the algebra of rationals.

‡ See E. V. Huntington, these Transactions, vol. 4 (1903), p. 31.

§ Schröder's notation will be employed throughout, except that x' (instead of x_1) will be used for the negative of x.

$$x\Delta y = axy + bxy' + cx'y + dx'y', \tag{1}$$

$$x \bigcirc y = pxy + qxy' + rx'y + sx'y'. \tag{2}$$

I proceed to find the conditions imposed by the postulates P_1–P_7 on the coefficients of (1) and(2).

The conditions that (1), (2) satisfy $P_1, P_2, P_3, P_4, P_5, P_6$ are* respectively

$$c = b, \tag{3}$$

$$a'd + (ad + a'd')(bc' + b'c) = 0, \tag{4}$$

$$b = a', d = c', \tag{5}$$

$$r = q, \tag{6}$$

$$p's + (ps + p's')(qr' + q'r) = 0, \tag{7}$$

$$a'(pq + rs) + d(p'q' + r's') + (ad + a'd' + bc + b'c')(p'q + r's) = 0. \tag{8}$$

To find the condition that (1), (2) satisfy P_7, note first that the condition $x\Delta x \neq x$ is the same as $a'x + dx' \neq 0$. Further, the condition that for given x, y there be an element η such that $x \bigcirc \eta = y$, is the condition of solvability in η of the equation

$$px\eta + qx\eta' + rx'\eta + sx'\eta' = y,$$

i. e. of the equation

$$(p'xy + pxy' + r'x'y' + rx'y')\eta + (q'xy + qxy' + s'x'y + sx'y')\eta' = 0.$$

The condition for the solvability of this equation is

$$p'q'xy + pqxy' + r's'x'y + rsx'y' = 0.$$

Hence the condition that (1), (2) satisfy P_7 is

$$p'q'xy + pqxy' + r's'x'y + rsx'y' = 0, a'x + dx' \neq 0, a'y + dy' \neq 0. \tag{9}$$

By (3)–(5), the condition that (1) satisfy P_1–P_3 simultaneously is

$$b = c = d' = a'.\dagger \tag{10}$$

By (6), (7), the condition that (2) satisfy P_4, P_5 simultaneously is

$$r = q, p's = 0. \tag{11}$$

By (8), (10), (11), the conditions that (1), (2) satisfy P_1–P_6 simultaneously are

---

* For (4), (5) see my *Operations with respect to which the elements of a boolean algebra form a group*, these Transactions, vol. 26 (1924), pp. 171–175, and E. Schröder, *Vorlesungen über die Algebra der Logik*, vol. 2, part 2, pp. 493, 494. My results were obtained without the knowledge of Schröder's work. Condition (8) I take from Schröder.

† So that all operations (1) with respect to which the elements of a boolean algebra form an abelian group, are given by $axy + a'xy' + a'x'y + ax'y'$. Compare my paper cited above.

$$b = c = d' = a', r = q, a'q + p'q + q's + ap'q' + aq's' = 0,$$

which reduce to

$$(12) \qquad b = c = d' = a', r = q = a, ap' + a's = 0.$$

By (10), if (1) satisfy P_1–P_3 then $d=a$, and the condition $a'x+dx'\neq 0$ in (9) becomes $a'x+ax'\neq 0$, i. e. $x\neq a$. It is to be observed that (9), (12) are sufficient as well as necessary conditions that (1), (2) satisfy P_1–P_7. Hence

THEOREM 1. *The necessary and sufficient conditions that* (1), (2) *satisfy* P_1-P_7 *simultaneously are*

$$(13) \qquad x\Delta y = axy + a'xy' + a'x'y + ax'y',$$

$$(14) \qquad x\bigcirc y = pxy + axy' + ax'y + sx'y',$$

$$(15) \qquad ap' + a's = 0,$$

$$(16) \qquad a'p'xy + apxy' + a's'x'y + asx'y' = 0, x \neq a, y \neq a.$$

Existence of fields. Conditions (13)-(16) do not, of course, prove that fields in boolean algebras exist. To consider existence, take first the algebra consisting of the boolean elements 0, 1. Let $a=0$ in (13)-(16). Then from (15), $s=0$. Letting $x=1$, $y=1$ in (16), we get $p=1$. Hence, *the boolean elements* 0, 1 *are a field for the pair of operations*

$$(17) \qquad xy' + x'y, \ xy.$$

Let $a=1$ in (13)-(16). Then from (15), $p=1$. From (16), by letting $x=0$, $y=0$, we get $s=0$. So that *the boolean elements* 0,1 *are a field for the pair of operations*

$$(18) \qquad xy + x'y', \ xy + xy' + x'y = x + y.$$

Since in a two-element boolean algebra all binary operations satisfying the condition of closure are of the form (1),* we observe from (13) that *operations* (17), (18) *are the only pair of operations for which the boolean elements* 0, 1 *are a field.*

Turning to a boolean algebra of more than two elements, let e be an element different from 0, 1. The element a in (13)–(16) must be 0, 1, or an element different from 0, 1. If $a=0$, then $s=0$. But by letting $x=e'$, $y=e$ in (16), we get $e=0$, contrary to hypothesis. We obtain a similar contradiction for $a=1$. Let a be different from 0,1. From (16), making the substitutions

$$x = 1, y = 1; x = 1, y = 0; x = 0, y = 1 \ ; x = 0, y = 0,$$

* See my *Complete sets of representations of two-element algebras*, Bulletin of the American Mathematical Society, vol. 30 (1924), p. 26. This paper also contains a different proof of the fact that (17), (18) are field operations for 0, 1.

we get respectively

$$a'p' = 0,\ ap = 0,\ a's' = 0,\ as = 0.$$

Hence

$$s = p = a'.$$

But this contradicts (15). We therefore find that *in a boolean algebra having more than two elements there exist no operations* (1), (2) *for which the elements are a field.*

To sum up, we have

THEOREM 2. *In a two-element boolean algebra there exist two and only two pairs of operations for which the elements are a field, namely*

$$xy' + x'y,\ xy;\quad xy + x'y',\ x + y.$$

In a boolean algebra of more than two elements there exist no operations expressible in terms of addition, multiplication, and negation for which the elements are a field.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

ASYMMETRIC DISPLACEMENT OF A VECTOR*

BY

JOSEPH MILLER THOMAS†

1. **Introduction.** Levi-Civita's definition of parallel displacement of a vector ‡ has been generalized to non-Riemannian geometries by several writers§ who have replaced the Christoffel symbols of the second kind by a set of quantities Γ^i_{jk} which are symmetric in j and k; and by Schouten|| who has omitted the assumption of symmetry.

The problem of the changes of the connection Γ^i_{jk} which preserve the paths (or geodesics) has been treated by Weyl, by Eisenhart and by Veblen in the symmetric case¶ and by Friedmann and Schouten** in the asymmetric case. Such changes of the connection preserve, in general, only the directions of vectors displaced along themselves. In the present paper are treated changes of connection which preserve the directions of *all* displaced vectors (§5).†† It is readily shown that two distinct symmetric connections cannot give rise to the same displaced directions, so that the connections considered are necessarily asymmetric in general.

In § 7 are given some tensors which are independent of the above change of connection, and a process, like covariant differentiation, for forming tensors of the same nature but of higher rank is indicated.

In the final section we find necessary and sufficient conditions in order that an asymmetric connection may be made symmetric by a change preserving displaced directions.

2. **General linear displacement of a vector.** Consider a vector field ξ^i in an n-dimensional manifold referred to a coördinate system x. The vector

* Presented to the Society, May 2, 1925, received by the editors in June, 1925.

† National Research Fellow in Mathematics.

‡ T. Levi-Civita, *Nozione di parallelismo in una varietà qualunque*, Rendiconti del Circolo Matematico di Palermo, vol. 42 (1917), pp. 173–205.

§ Cf. H. Weyl, *Raum, Zeit, Materie*, 4th edition, p. 100; A. S. Eddington, *The Mathematical Theory of Relativity*, p. 213.

|| J. A. Schouten, *Über die verschiedenen Arten der Übertragung*, Mathematische Zeitschrift, vol. 13 (1922), pp. 56–81; *Der Ricci-Kalkül*, pp. 62–75.

¶ H. Weyl, Göttinger Nachrichten, 1921, p. 99; L. P. Eisenhart, Proceedings of the National Academy of Sciences, vol. 8 (1922), p. 233; O. Veblen, ibid., p. 347.

** Mathematische Zeitschrift, vol. 21 (1924), p. 218.

†† This problem has been touched upon by H. Friesecke in his paper *Vektorübertragung, Richtungsübertragung, Metrik*, Mathematische Annalen, vol. 94 (1925), p. 101.

ξ^i at a point P of coördinates x^i can be thought of as being displaced to a nearby point P' of coördinates x^i+dx^i provided there is given a law specifying the vector at P' which corresponds to the vector ξ^i at P, or, what amounts to the same thing, a law specifying the vector $\delta\xi^i$ at P' which is the difference between the vector of the field at the point P' and the displaced vector. Adopting the latter alternative,* we say that a vector ξ^i suffers a *general linear displacement* when

$$\delta\xi^i = d\xi^i + H^i_{jk}\xi^j dx^k, \tag{2.1}$$

where $d\xi^i$ represent the differentials of the functions ξ^i.

The quantities H^i_{jk}, which will be called the components of the linear connection, are subjected only to the restriction that the quantities given by (2.1) are the components of a vector—a restriction which determines their law of transformation. When the coördinates x are changed by an analytic transformation to $\bar{x}$, we have

$$d\bar{\xi}^i = \frac{\partial\bar{x}^i}{\partial x^\alpha} d\xi^\alpha + \frac{\partial^2\bar{x}^i}{\partial x^j\partial x^k}\xi^j dx^k,$$

$$\bar{H}^i_{jk}\bar{\xi}^j d\bar{x}^k = \bar{H}^i_{\alpha\beta}\frac{\partial\bar{x}^\alpha}{\partial x^j}\frac{\partial\bar{x}^\beta}{\partial x^k}\xi^j dx^k.$$

Hence

$$\delta\bar{\xi}^i = \frac{\partial\bar{x}^i}{\partial x^\alpha} d\xi^\alpha + \left(\frac{\partial^2\bar{x}^i}{\partial x^j\partial x^k} + \bar{H}^i_{\alpha\beta}\frac{\partial\bar{x}^\alpha}{\partial x^j}\frac{\partial\bar{x}^\beta}{\partial x^k}\right)\xi^j dx^k.$$

In order that these equations may reduce to

$$\frac{\partial\bar{x}^i}{\partial x^\alpha}\delta\xi^\alpha = \frac{\partial\bar{x}^i}{\partial x^\alpha} d\xi^\alpha + \frac{\partial\bar{x}^i}{\partial x^\alpha} H^\alpha_{jk}\xi^j dx^k,$$

it is necessary and sufficient that

$$H^\alpha_{jk}\frac{\partial\bar{x}^i}{\partial x^\alpha} = \frac{\partial^2\bar{x}^i}{\partial x^j\partial x^k} + \bar{H}^i_{\alpha\beta}\frac{\partial\bar{x}^\alpha}{\partial x^j}\frac{\partial\bar{x}^\beta}{\partial x^k}. \tag{2.2}$$

We write

$$H^i_{jk} = \Gamma^i_{jk} + \Omega^i_{jk}, \tag{2.3}$$

---

* Although displacement as here defined only applies to a vector belonging to a field, we extend the definition to all vectors by stipulating that the changes in the components be given by $-H^i_{jk}\xi^j dx^k$ as far as terms of the first order.

the Γ's denoting the symmetric part of H^i_{jk},

$$\Gamma^i_{jk} = \tfrac{1}{2}\,(H^i_{jk} + H^i_{kj}), \tag{2.4}$$

and the Ω's the skew-symmetric part,

$$\Omega^i_{jk} = \tfrac{1}{2}\,(H^i_{jk} - H^i_{kj}). \tag{2.5}$$

By adding equations (2.2) to those obtained from them by the interchange of j and k, and dividing by 2, we find the law of transformation for the Γ's is the same as for the H's, namely,

$$\Gamma^\alpha_{jk}\frac{\partial \bar{x}^i}{\partial x^\alpha} = \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} + \bar{\Gamma}^i_{\alpha\beta}\frac{\partial \bar{x}^\alpha}{\partial x^j}\,\frac{\partial \bar{x}^\beta}{\partial x^k}\,. \tag{2.6}$$

From these equations it is readily proved that

$$\Gamma^\alpha_{\alpha j} = \bar{\Gamma}^\alpha_{\alpha\beta}\frac{\partial \bar{x}^\beta}{\partial x^j} + \frac{\partial \log \Delta}{\partial x^j}, \tag{2.7}$$

where

$$\Delta = \left|\,\frac{\partial \bar{x}^i}{\partial x^j}\,\right|\,.$$

By subtracting equations of the form (2.2), we prove that Ω^i_{jk} are the components of a tensor.*

3. **Fields of vectors generated by displacement and the curvature tensor.** Let us consider a field of vectors generated by the displacement of a vector. The conditions for such a field are $\delta\xi^i = 0$, and can be written as

$$d\xi^i + H^i_{jk}\xi^j dx^k = 0, \tag{3.1}$$

or as

$$\frac{\partial \xi^i}{\partial x^k} + H^i_{jk}\xi^j = 0. \tag{3.2}$$

The conditions of integrability of these equations are

$$Z^i_{jkl}\xi^j = 0, \tag{3.3}$$

where

$$Z^i_{jkl} = \frac{\partial H^i_{jk}}{\partial x^l} - \frac{\partial H^i_{jl}}{\partial x^k} + H^\alpha_{jk}H^i_{\alpha l} - H^\alpha_{jl}H^i_{\alpha k}. \tag{3.4}$$

---

* This result is given by Schouten, *Ricci-Kalkül*, p. 67.

In order that there exist a vector at each point of space which can be obtained by displacement of a vector which has arbitrary initial components at an arbitrary starting point, it is necessary and sufficient that equations (3.2) be completely integrable. The conditions of integrability (3.3) of these equations will be satisfied identically if, and only if,

$$Z^i_{jkl} = 0.$$

From the way in which these functions Z^i_{jkl} arise it is evident that they are the components of a tensor. By substitution from (2.3) we have

$$(3.5) \qquad Z^i_{jkl} = B^i_{jkl} + C^i_{jkl},$$

where $B^i_{jk,l}$ is the ordinary curvature tensor for the Γ's, namely,

$$(3.6) \qquad B^i_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial x^l} - \frac{\partial \Gamma^i_{jl}}{\partial x^k} + \Gamma^\alpha_{jk}\Gamma^i_{\alpha l} - \Gamma^\alpha_{jl}\Gamma^i_{\alpha k},$$

and where

$$(3.7) \qquad C^i_{jkl} = \Omega^i_{jk,l} - \Omega^i_{jl,k} + \Omega^\alpha_{jk}\Omega^i_{\alpha l} - \Omega^\alpha_{jl}\Omega^i_{\alpha k},$$

$\Omega^i_{jk,l}$ being the covariant derivative of Ω^i_{jk} formed with respect to the Γ's:

$$(3.8) \qquad \Omega^i_{jk,l} = \frac{\partial \Omega^i_{jk}}{\partial x^l} + \Omega^\alpha_{jk}\Gamma^i_{\alpha l} - \Omega^i_{\alpha k}\Gamma^\alpha_{jl} - \Omega^i_{j\alpha}\Gamma^\alpha_{kl}.$$

4. **The paths or geodesics.** If we define a path (geodesic) as a curve whose tangent vector at any point is obtained by displacement of the tangent vector at a nearby point, the differential equations of the paths are

$$(4.1) \qquad \frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0.$$

Only the symmetric part of the coefficients H^i_{jk} enters into these equations.*

By choosing for Ω^i_{jk} an arbitrary tensor, skew-symmetric in j and k, we can associate with any geometry of paths† a displacement of the type (2.1).

5. **Changes of linear connection which preserve displaced directions.** The vector arising from ξ^i by displacement to the point x^i+dx^i has for its components, as far as terms of the first order,

$$(5.1) \qquad \xi^i - H^i_{jk}\xi^j dx^k.$$

* Schouten, *Ricci-Kalkül*, p. 76.

† Cf. O. Veblen and T. Y. Thomas, *The geometry of paths*, these Transactions, vol. 25 (1923), p. 551–580, for a general account of the geometry of paths.

If we consider another set of H's, say

$$H''^{i}_{jk} = H^{i}_{jk} - a^{i}_{jk}, \tag{5.2}$$

then the new vector arising by displacement has for its components

$$\xi^i - H^{i}_{jk}\xi^j dx^k + a^{i}_{jk}\xi^j dx^k. \tag{5.3}$$

In order that the directions of the vectors (5.1) and (5.3) may be the same, it is necessary and sufficient that their difference be in the direction of either, or that

$$a^{i}_{jk}\xi^j dx^k = \lambda\xi^i, \tag{5.4}$$

to within terms of the first order. The conditions can also be written

$$a^{i}_{jk}\xi^j dx^k = \lambda\delta^{i}_{j}\xi^j.$$

Since these equations must hold for arbitrary ξ^j, we have

$$a^{i}_{jk}dx^k = \lambda\delta^{i}_{j}.$$

If we determine λ from this last relation by contraction, substitute back and equate the coefficients of dx^k to zero, we find

$$a^{i}_{jk} = \frac{\delta^{i}_{j}}{n} a^{\alpha}_{\alpha k}. \tag{5.5}$$

It is seen from (2.2) and (5.2) that a^{i}_{jk} is a tensor. Hence $a^{\alpha}_{\alpha k}$ is a vector, which will be denoted by $2n\varphi_k$. In terms of this vector, equations (5.5) become*

$$a^{i}_{jk} = 2\delta^{i}_{j}\varphi_k,$$

and from (2.4) and (2.5) the changes in the Γ's and Ω's are, therefore, respectively

$$\Gamma^{i}_{jk} - \Gamma''^{i}_{jk} = \delta^{i}_{j}\varphi_k + \delta^{i}_{k}\varphi_j, \tag{5.6}$$

$$\Omega^{i}_{jk} - \Omega''^{i}_{jk} = \delta^{i}_{j}\varphi_k - \delta^{i}_{k}\varphi_j. \tag{5.7}$$

It is to be noted that the skew-symmetric part of the change vanishes only if the symmetric part does and vice-versa.

Conversely, if the connection is changed in accordance with (5.6) and (5.7), φ_j being an arbitrary vector, conditions (5.4) are fulfilled.

In the associated geometry of paths, there is brought about a change of the affine connection Γ^{i}_{jk} which preserves the paths. Moreover, every

* Friesecke, loc. cit., p. 106, obtains these conditions.

change preserving the paths is of the form (5.6).* Hence a study of displacements of the form (2.1) is one approach to a *projective* geometry of paths.

From (5.6) and (5.7) follow

$$\varphi_j = \frac{1}{n+1}\left(\Gamma^\alpha_{\alpha j} - \Gamma'^\alpha_{\alpha j}\right), \tag{5.8}$$

$$\varphi_j = \frac{1}{n-1}\left(\Omega'^\alpha_{j\alpha} - \Omega^\alpha_{j\alpha}\right). \tag{5.9}$$

From these equations, (5.6), and (5.7) we find that the following expressions are independent of φ_j and are therefore invariant under the change of connection being considered:†

$$\Pi^i_{jk} = \Gamma^i_{jk} - \frac{\delta^i_j}{n+1}\Gamma^\alpha_{\alpha k} - \frac{\delta^i_k}{n+1}\Gamma^\alpha_{\alpha j}\ ; \tag{5.10}$$

$$\mathfrak{L}^i_{jk} = \Omega^i_{jk} - \frac{\delta^i_j}{n-1}\Omega^\alpha_{\alpha k} - \frac{\delta^i_k}{n-1}\Omega^\alpha_{j\alpha}\ ; \tag{5.11}$$

$$\Sigma^i_{jk} = \Gamma^i_{jk} + \frac{\delta^i_j}{n-1}\Omega^\alpha_{k\alpha} + \frac{\delta^i_k}{n-1}\Omega^\alpha_{j\alpha}\ ; \tag{5.12}$$

$$\frac{1}{n+1}\Sigma^\alpha_{\alpha k} = \frac{\Gamma^\alpha_{\alpha k}}{n+1} + \frac{\Omega^\alpha_{k\alpha}}{n-1}\ ; \tag{5.13}$$

$$L^i_{jk} = H^i_{jk} - \frac{\delta^i_j}{n}H^\alpha_{\alpha k}. \tag{5.14}$$

We note then that from these definitions, it follows that

$$\Pi^\alpha_{\alpha k} = \mathfrak{L}^\alpha_{\alpha k} = L^\alpha_{\alpha k} = 0.$$

6. **Normal affine connection.** The quantities Π^i_{jk} defined by (5.10) are the components of the projective connection.‡ From (5.10) it follows that

$$\Pi^i_{jk} = \Gamma^i_{jk} \tag{6.1}$$

---

* Cf. H. Weyl, Göttinger Nachrichten, 1921, p. 99. Also L. P. Eisenhart, Proceedings of the National Academy of Sciences, vol. 8 (1922), p. 233, and O. Veblen, ibid., p. 347.

† We shall use German characters to denote tensors which are independent of φ.

‡ These quantities were first employed and named by T. Y. Thomas in the Proceedings of the National Academy of Sciences, vol. 11 (1925), pp. 199–203.

if, and only if, $\Gamma^{\alpha}_{\alpha i}=0$. From (2.7) it is seen that this will be an invariant property only under transformations of constant jacobian. An affine connection for which (6.1) is true will be called normal* for the given coördinate system and those arising from it by transformations with constant jacobian.†

If *in the given coördinate system* we choose the components φ_i equal to $\Gamma^{\alpha}_{\alpha i}/(n+1)$, then from (5.8) it follows that

$$\Gamma'^{\alpha}_{\alpha i} = 0, \tag{6.2}$$

and conversely. Hence we have proved

THEOREM 1. *For any coördinate system there is a unique normal affine connection.*

T. Y. Thomas‡ states that the components of the projective connection "constitute a normalized affine connection." In the light of the above discussion this statement must be interpreted as follows: for each coördinate system there exists an affine connection whose components are equal to the corresponding components of the projective connection in the given coördinate system. The same equality of components also holds in coördinate systems arising from the given one by transformations of constant jacobian. In coördinate systems connected with the given one by transformations of variable jacobian, however, a *different* affine connection will be normalized and will have its components equal to those of the projective connection.

It also follows that the skew-symmetric tensor

$$S_{ij} = B^{\alpha}_{\alpha ij} = \frac{\partial \Gamma^{\alpha}_{\alpha i}}{\partial x^{j}} - \frac{\partial \Gamma^{\alpha}_{\alpha j}}{\partial x^{i}} \tag{6.3}$$

vanishes for a normal affine connection in a properly chosen coördinate system, and therefore in any coördinate system.§ Conversely, we can show

---

* E. Cartan, adopting a viewpoint entirely different from that of the present paper, includes $\Gamma^{\alpha}_{\alpha i} = 0$ in the definition of his normal projective connection. He proves a result equivalent to Theorem 1. Cf. his paper *Sur les variétés à connexion projective*, Bulletin de la Société Mathématique de France, vol. 52 (1924), p. 211.

† T. Y. Thomas (loc. cit.) places $\Delta = 1$ as a device to make the Π's have a law of transformation like that of the Γ's. He calls transformations of coördinates of this type *equi-transformations*. For a determination in finite form of the equations of all such transformations, cf. E. Goursat, Bulletin des Sciences Mathématiques, vol. 41 (1917), p. 211.

‡ Loc. cit., p. 200.

§ That it is always possible to choose the affine connection so that the skew-symmetric tensor vanishes was first established by L. P. Eisenhart, Proceedings of the National Academy of Sciences, vol. 8 (1922), p. 233.

that if the skew-symmetric tensor vanishes, then it is possible to choose the coördinate system so that $\bar{\Gamma}^{\alpha}_{\alpha i}=0$. We have by hypothesis

$$\frac{\partial \Gamma^{\alpha}_{\alpha i}}{\partial x^{j}} = \frac{\partial \Gamma^{\alpha}_{\alpha j}}{\partial x^{i}}.$$

Hence the equations

$$\frac{\partial \log \Delta}{\partial x^{i}} = \Gamma^{\alpha}_{\alpha i}$$

are completely integrable, and it is possible to determine a Δ satisfying them and assuming an initial value different from 0. Suppose that such a solution is

$$\Delta = f(x^1, x^2, \cdots, x^n). \tag{6.4}$$

Choose the variables $(\bar{x}^2, \bar{x}^3, \cdots, \bar{x}^n)$ as arbitrary analytic functions of $(x^1, x^2, \cdots, x^n)$ subject only to the restriction that the minor of $\partial\bar{x}^1/\partial x^1$ in Δ shall initially be different from zero. Then (6.4) can be solved for this derivative and Cauchy's theorem applied to show the existence of functions $\bar{x}^1$ satisfying (6.4). In the coördinate system so obtained it follows from (2.7) that $\bar{\Gamma}^{\alpha}_{\alpha i}=0$.

THEOREM 2. *A necessary and sufficient condition for the existence of coördinate systems for which a given affine connection is normal is the vanishing of the skew-symmetric tensor.*

Veblen* has pointed out that for affine connections of the type under consideration there is a definition of volume:

$$V = \int \gamma dx^1 dx^2 \cdots dx^n, \qquad \frac{\partial \log \gamma}{\partial x^{i}} = \Gamma^{\alpha}_{\alpha i}.$$

In the coördinate systems referred to in Theorem 2 this definition takes the form

$$V = \gamma \int dx^1 dx^2 \cdots dx^n,$$

γ being a constant. In a Riemann geometry, the skew-symmetric tensor is zero. Hence coördinate systems exist for which the affine connection given by the Christoffel symbols is normal. Since in this case

$$\Gamma^{\alpha}_{\alpha i} = \frac{\partial \log \sqrt{g}}{\partial x^{i}},$$

* O. Veblen, Proceedings of the National Academy of Sciences, vol. 9 (1923), p. 3. Cf. also L. P. Eisenhart, ibid., p. 4.

where g is the determinant of the fundamental tensor g_{ij}, they are the coördinate systems for which g is constant.

It is likewise seen from (5.9) that by choosing φ_i properly we can make

$$\Omega^{\alpha}_{\alpha j} = 0$$

for the displacement, so that

$$\mathfrak{L}^{i}_{jk} = \Omega^{i}_{jk}. \tag{6.5}$$

These are invariant conditions, but in general they can not be realized simultaneously with (6.1).

A proper choice of φ_i will in a similar manner make

$$L^{i}_{jk} = H^{i}_{jk} \tag{6.6}$$

for the given coördinate system.

7. Tensors independent of the change of the linear connection. By computation from (3.6) or by reference to Veblen and T. Y. Thomas* we find

$$B'^{i}_{jkl} = B^{i}_{jkl} - \delta^{i}_{j}(\Phi_{kl} - \Phi_{lk}) - \delta^{i}_{k}\Phi_{jl} + \delta^{i}_{l}\Phi_{jk}, \tag{7.1}$$

where

$$\Phi_{kl} = \varphi_{k,l} + \varphi_k\varphi_l,$$

and $\varphi_{k,l}$ is the covariant derivative of φ_k formed with respect to the Γ's. By contraction, the following expressions are found:

$$\Phi_{jk} = \frac{R'_{jk} - R_{jk}}{n-1} + \frac{S'_{jk} - S_{jk}}{n^2 - 1}, \tag{7.2}$$

$$\Phi_{jk} - \Phi_{kj} = \frac{S_{jk} - S'_{jk}}{n+1}, \tag{7.3}$$

where $R_{jk} = B^{\alpha}_{jk\alpha}$ is the Ricci tensor and S_{ij} is given by (6.3). Substitution from (7.2) in (7.1) and separation of the accented and unaccented terms show that the quantities

$$\begin{aligned}\mathfrak{W}^{i}_{jkl} = B^{i}_{jkl} - \frac{\delta^{i}_{j}}{n+1}S_{kl} &+ \frac{1}{n-1}(\delta^{i}_{k}R_{jl} - \delta^{i}_{l}R_{jk}) \\ &+ \frac{1}{n^2-1}(\delta^{i}_{k}S_{jl} - \delta^{i}_{l}S_{jk})\end{aligned} \tag{7.4}$$

---

* Loc. cit., p. 559; also L. P. Eisenhart, Annals of Mathematics, ser. 2, vol. 24 (1923), p. 377, where there are differences of sign due to different definitions of φ_i and B^{i}_{jkl}.

are the components of a tensor which is independent of the vector φ. It is the projective curvature tensor discovered by Weyl.* It can also be obtained by expressing integrability conditions of the equations of transformation of the components of the projective connection (5.10).† When it is obtained in the latter way, it is expressed in terms of the Π's. That it can be so expressed, follows directly from the observations leading to (6.1). In fact, we can prove

THEOREM 3. *Any projective tensor formed from the components of the affine connection can be expressed in terms of the components of the projective connection.*

By a *projective tensor* is meant a tensor independent of the vector φ. To prove the theorem, we compute the values of the components of the tensor for a normal affine connection. Equations (6.1) show that each Γ can be replaced by the corresponding Π. Since the values of the components of the tensor are by hypothesis independent of the affine connection for which they are computed, the theorem is proved.

Treating (3.7) in an entirely analogous manner, we get

$$C''^{i}_{jkl} = C^{i}_{jkl} - \delta^{i}_{j}(\Phi_{kl} - \Phi_{lk}) + \delta^{i}_{k}\Phi_{jl} - \delta^{i}_{l}\Phi_{jk},$$

$$\Phi_{jk} - \Phi_{kj} = \frac{D_{jk} - D'_{jk}}{n-1},$$

$$\Phi_{jk} = \frac{C_{jk} - C'_{jk}}{n-1} + \frac{D_{jk} - D'_{jk}}{(n-1)^2},$$

$$C_{jk} = C^{\alpha}_{jk\alpha}, \quad D_{jk} = C^{\alpha}_{\alpha jk}.$$

It then is found that the following are the components of tensors independent of φ:

$$\mathfrak{X}^{i}_{jkl} = C^{i}_{jkl} - \frac{\delta^{i}_{j} D_{kl}}{n-1} + \frac{1}{n-1}(\delta^{i}_{k}C_{jl} - \delta^{i}_{l}C_{jk}) + \frac{1}{(n-1)^2}(\delta^{i}_{k}D_{jl} - \delta^{i}_{l}D_{jk}), \tag{7.5}$$

* H. Weyl, Göttinger Nachrichten, 1921, p. 99.

† Cf. a paper by the writer, Proceedings of the National Academy of Sciences, vol. 11 (1925), p. 207.

$$\mathfrak{X}^{\alpha}_{\alpha jk} = \frac{(3-n)D_{jk}}{(n-1)^2} + \frac{C_{jk} - C_{kj}}{n-1}, \tag{7.6}$$

$$\mathfrak{Y}_{jk} = \frac{D_{jk}}{n-1} - \frac{S_{jk}}{n+1}. \tag{7.7}$$

The tensor given by (7.7) can also be obtained by expressing integrability conditions of the equations of transformation of the quantities (5.13).

We find also

$$\mathfrak{W}^{\alpha}_{\alpha kl} = \mathfrak{W}^{\alpha}_{kl\alpha} = \mathfrak{X}^{\alpha}_{kl\alpha} = 0,$$

and that the Weyl tensor has the property of cyclic symmetry, namely,

$$\mathfrak{W}^{i}_{jkl} + \mathfrak{W}^{i}_{klj} + \mathfrak{W}^{i}_{ljk} = 0,$$

whereas the tensor $\mathfrak{X}^{i}_{jkl}$ does not.

We note also that the tensor

$$\mathfrak{Z}^{i}_{jkl} = Z^{i}_{jkl} - \frac{\delta^{i}_{j}}{n} Z^{\alpha}_{\alpha kl}$$

arises by treating (3.4) in the same way that we did (3.6) and (3.7). From a consideration of equations (6.6) we see that this tensor can be expressed in terms of the quantities (5.14), and a theorem similar to Theorem 3 can be stated for projective tensors formed from the H's.

In addition to the above tensors which are independent of φ, we have also the projective invariants given by formulas (5.10) to (5.14). Of these, the quantities Σ^{i}_{jk} present special interest. They constitute an affine (symmetric) connection which is uniquely determined by the linear connection and which is the same for all linear connections yielding the same displaced directions. Since it is readily proved that the Σ's have the same law of transformation as the Γ's, they form a basis for covariant differentiation which is independent of φ. Thus from the projective tensors given in this section we can form infinite sequences of tensors of the same character. To obtain formulas for covariant differentiation we need only replace the Γ's by Σ's in the ordinary formulas.*

8. **Semi-symmetric displacements.** We next inquire under what conditions it is possible to choose φ so that $\Omega^{i}_{jk}=0$. A necessary condition is obviously the vanishing of the tensor $\mathfrak{Q}^{i}_{jk}$, that is,

$$\Omega^{i}_{jk} = \frac{\delta^{i}_{j}}{n-1}\Omega^{\alpha}_{\alpha k} + \frac{\delta^{i}_{k}}{n-1}\Omega^{\alpha}_{j\alpha}. \tag{8.1}$$

* Cf. O. Veblen and T. Y. Thomas, loc. cit., p. 571.

This condition is also sufficient; for, as remarked above, the vector φ can be chosen so as to reduce $\mathfrak{L}^{i}_{jk}$ to Ω^{i}_{jk}. We can write (8.1) in the form

$$\Omega^{i}_{jk} = \delta^{i}_{j}\psi_{k} - \delta^{i}_{k}\psi_{j}, \tag{8.2}$$

where ψ_j is a vector. This gives the type of asymmetric displacement studied by Friedmann and Schouten and called by them a semi-symmetric displacement (halb-symmetrische Übertragung).* We can therefore state the following theorems.

THEOREM 4. *A necessary and sufficient condition that an asymmetric displacement be semi-symmetric is the vanishing of the tensor* $\mathfrak{L}^{i}_{jk}$.

THEOREM 5. *A semi-symmetric displacement can always be replaced by a symmetric one with preservation of displaced directions.*

It is seen from (5.12) that for this displacement, reduced to the symmetric form,

$$\Sigma^{i}_{jk} = \Gamma^{i}_{jk}.$$

We can also arrive at the semi-symmetric displacement in another manner. Let the components of the linear connection in a coördinate system y be denoted by $\overline{H}^{i}_{jk}$, and consider a vector with components η^i in that system. The vector η^i and that obtained from it by displacement, namely

$$\eta^i - \overline{H}^{i}_{jk}\eta^j dy^k,$$

will have proportional components in the y coördinate system, if, and only if,

$$\overline{H}^{i}_{jk} = \frac{\delta^{i}_{j}}{n}\overline{H}^{\alpha}_{\alpha k} \tag{8.3}$$

at the given point. The proof of the foregoing statement is entirely analogous to the derivation of equations (5.5). Denoting by p_k the expressions

$$\frac{1}{n}\overline{H}^{\alpha}_{\alpha\beta}\frac{\partial y^\beta}{\partial x^k}$$

we find from (2.2)

$$H^{i}_{jk} = \frac{\partial^2 y^\alpha}{\partial x^j \partial x^k}\frac{\partial x^i}{\partial y^\alpha} + \delta^{i}_{j}p_k,$$

---

* Loc. cit. Cf. also J. A. Schouten, Proceedings, Koninklijke Akademie van Wetenschappen te Amsterdam, vol. 26 (1923), p. 850, where certain applications to physics are indicated. In the latter connection, cf. H. Eyraud, Comptes Rendus, January, 1925, pp. 127–129.

whence by interchange of j and k and subtraction,

$$2\Omega^i_{jk} = \delta^i_j p_k - \delta^i_k p_j. \tag{8.4}$$

Contraction gives

$$p_k = \frac{2}{n-1}\Omega^\alpha_{\alpha k},$$

and substitution in (8.4) shows that (8.1) are satisfied. Hence the displacement is semi-symmetric. Conversely, if (8.1) are fulfilled and we define a system of coördinates with origin at the point $x^i = x^i_0$ by the equations

$$x^i = x^i_0 + y^i - \tfrac{1}{2}(\Sigma^i_{\alpha\beta})_0 y^\alpha y^\beta, \tag{8.5}$$

then by means of equations (2.2) written in the form

$$\bar{H}^i_{jk} = \frac{\partial^2 x^\alpha}{\partial y^j \partial y^k}\frac{\partial y^i}{\partial x^\alpha} + H^\alpha_{\beta\gamma}\frac{\partial y^i}{\partial x}\frac{\partial x^\beta}{\partial y^j}\frac{\partial x^\gamma}{\partial y^k},$$

and the relations

$$\left(\frac{\partial x^i}{\partial y^j}\right)_0 = \delta^i_j, \quad \left(\frac{\partial^2 x^i}{\partial y^j \partial y^k}\right)_0 = -(\Sigma^i_{jk})_0,$$

which are consequences of (8.5), we find that

$$(\bar{H}^i_{jk})_0 = \frac{2\delta^i_j(\Omega^\alpha_{\alpha k})_0}{n-1}.$$

Hence (8.3) are satisfied, and it is seen that the quantities (5.14) vanish at the origin in the y coördinate system.

THEOREM 6. *A necessary and sufficient condition in order that a displacement be semi-symmetric is that there exist for each point of space a coördinate system in terms of which any vector at the given point and that arising from it by displacement to a nearby point have proportional components.**

It is to be noted that the coördinate system given by (8.5) is independent of the vector φ and is the geodesic coördinate system for the associated symmetric connection.

I wish to express my thanks to Professors Eisenhart and Veblen, who have read the manuscript of this paper and improved it by their suggestions.

* Weyl, in *Raum, Zeit, Materie*, p. 100, arrives at the affine connection by postulating the existence of a coördinate system in terms of which original and displaced vectors have equal components.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

TENSORS DETERMINED BY A HYPERSURFACE IN RIEMANN SPACE*

BY

HARRY LEVY

Introduction

Bianchi† has generalized the study of a surface in euclidean 3-space by considering a hypersurface, V_n, immersed in an arbitrary enveloping space of one more dimension. Associated with such a hypersurface there are two differential quadratic forms, which in this paper we denote by $g_{\alpha\beta}du^\alpha du^\beta$ and $\Omega_{\alpha\beta}du^\alpha du^\beta$, where the u's $(\alpha, \beta = 1, 2, \cdots, n)$ are coördinates in V_n. These forms, known as the first and second fundamental forms of the hypersurface respectively, are defined precisely as are the two fundamental forms of a surface in ordinary 3-space; the first in the case of a positive definite form gives the square of the element of the arc,

$$ds^2 = g_{\alpha\beta}du^\alpha du^\beta,$$

while the second may be obtained by considering the variation of the first order of the fundamental tensor of the space as one passes from the given hypersurface to a nearby hypersurface geodesically parallel to the given one. The variations of higher order, not considered by Bianchi, yield additional tensors which are intimately connected with the study of the given hypersurface and of the hypersurfaces geodesically parallel to the given one. These tensors can be found by writing the linear element of the enveloping space in the form

$$\varphi = e(du^0)^2 + c_{\alpha\beta}du^\alpha du^\beta,$$

where the curves of parameter u^0 are geodesics orthogonal to V_n and the parameter u^0 is the arc of these geodesics measured from V_n. The coefficients in the expression of $c_{\alpha\beta}$ as a power series in u^0 are the required sequence of tensors,

$$c_{\alpha\beta} = g_{\alpha\beta} - 2\Omega_{\alpha\beta}u^0 + \cdots.$$

The hypersurfaces $u^0 = $ constant are geodesically parallel and the fundamental tensor of any one of them is given by this power series with u^0

---

* Presented to the Society, February 28, 1925; received by the editors in February, 1925.

† Bianchi, *Lezioni di Geometria Differenziale*, 2d edition, vol. I, or 3d edition, vol. II. Reference may also be made to other works that have recently appeared, *Der Ricci-Kalkül*, by J. A. Schouten; *Lezioni di Calcolo Differenziale Assoluto*, by T. Levi-Civita; *Vorlesungen über Differentialgeometrie*, II, by W. Blaschke; *Riemannian Geometry*, by L. P. Eisenhart, Princeton University Press.

replaced by the particular value of the constant. Hence these remaining terms are essential, in particular, in the study of geodesically parallel hypersurfaces.

We devote the first part of the paper to a brief capitulation of known results, and to the proof of two theorems in general tensor analysis which we need in the later development. In the second part we obtain the general form of this sequence of tensors, expressing it in intrinsic form in terms of known functions of the space and hypersurface. In the last part of the paper we discuss some subcases, proving several theorems of geometric interest.

The author wishes to express his appreciation of the valuable suggestions and helpful criticism given him by Professor Eisenhart.

Part I

1. We consider a Riemann space of $n+1$ dimensions, denoted for brevity by V_{n+1}; the fundamental form of this space*

$$\varphi = a_{rs}dx^r dx^s, \tag{1.1}$$

is assumed to have a non-vanishing determinant,

$$a = |a_{rs}| \neq 0. \tag{1.2}$$

We denote by a^{rs} the cofactor of a_{rs} in the determinant a divided by a itself. We may then write

$$a_{rs}a^{st} = \delta_r^t \qquad (r,t = 0,1,2,\cdots,n), \tag{1.3}$$

where δ_r^t is a Kronecker delta, that is,

$$\delta_r^t = \left.\begin{matrix} 0, & r \neq t \\ 1, & r = t \end{matrix}\right\} \qquad (r,t = 0,1,\cdots,n). \tag{1.4}$$

Let $[rs, t]$ and $\left\{ {p \atop rs} \right\}$ be the Christoffel symbols of the first and second kinds respectively, so that

$$[rs,t] = \frac{1}{2}\left(\frac{\partial a_{rt}}{\partial x^s} + \frac{\partial a_{st}}{\partial x^r} - \frac{\partial a_{rs}}{\partial x^t}\right) \tag{1.5}$$

and

$$\left\{ {p \atop r\,s} \right\} = a^{pt}[rs,t], \quad [rs,t] = a_{pt}\left\{ {p \atop r\,s} \right\}. \tag{1.6}$$

---

* As usual, the repetition of an index, once as subscript and once as superscript, indicates a summation on that index over all values from 0 to n. However, compare the remarks following equations (4.3).

We need also the Riemann symbols of the fourth order; denote those of the first kind by R_{prst} and those of the second kind by R^p_{rst}, so that

$$(1.7)\quad R_{prst} = \frac{\partial}{\partial x^t}[ps,r] - \frac{\partial}{\partial x^s}[pt,r] + \left\{ {k \atop pt} \right\}[rs,k] - \left\{ {k \atop ps} \right\}[rt,k],$$

$$(1.8)\quad R^r_{pst} = \frac{\partial}{\partial x^t}\left\{ {r \atop ps} \right\} - \frac{\partial}{\partial x^s}\left\{ {r \atop pt} \right\} + \left\{ {k \atop ps} \right\}\left\{ {r \atop kt} \right\} - \left\{ {k \atop pt} \right\}\left\{ {r \atop ks} \right\}$$

and

$$(1.9)\qquad R^p_{rst} = a^{pq}R_{rqst}, \qquad R_{prst} = a_{rq}R^q_{pst}.$$

These relations may be written in a number of other ways because of (1.5), (1.6) and (1.9).* The properties of the R's, namely that

$$(1.10)\qquad R^p_{rst} = -R^p_{rts},$$

$$(1.11)\qquad R^p_{rst} + R^p_{str} + R^p_{trs} = 0,$$

$$(1.12)\qquad R_{prst} = -R_{prts} = R_{rpts} = R_{tsrp},$$

follow quite readily from the definitions. The reader is referred to Bianchi for further details.

Let us denote the covariant derivative of R_{prst} with respect to x^e and the form (1.1) by $R_{prst,e}$ and similarly let $R^p_{rst,e}$ be the covariant derivative of R^p_{rst}. Since the covariant derivatives of a_{ij} and a^{ij} vanish,† it follows from (1.9) that

$$(1.13)\qquad R^p_{rst,e} = a^{pq}R_{rqst,e}, \qquad R_{prst,e} = a_{rq}R^q_{pst,e}.$$

Let us recall also the identity of Bianchi‡

$$(1.14)\qquad R^p_{rst,u} + R^p_{rtu,s} + R^p_{rus,t} = 0.$$

Successive covariant differentiation of $R^p_{rst,u}$ and $R_{prst,u}$ gives us a sequence of tensors which we denote by $R^p_{rst,u\cdots v}$ and $R_{prst,u\cdots v}$. We find that these R's satisfy the analogues of (1.13) and the equations obtained through the differentiation of (1.10), (1.11), (1.12) and (1.14).

---

* Bianchi, 2d edition, loc. cit., pp. 72, 73.

† Ricci and Levi-Civita, *Méthodes de calcul différential absolu*, Mathematische Annalen, vol. 54 (1901), p. 138.

‡ Bianchi, loc. cit., p. 351. Also Veblen, *Normal coördinates for the geometry of paths*, Proceedings of the National Academy of Sciences, vol. 8 (1922), p. 197.

2. Consider a point function $f(x^0, x^1, \cdots, x^n)$ evaluated along a curve C. If we write

$$\lambda^r = \frac{dx^r}{ds} \qquad (r = 0,1,2, \cdots, n), \tag{2.1}$$

the λ's are the components of the vector tangent to the curve, s being its arc;* the derivative of f along C is given by

$$\frac{\partial f}{\partial s} = f_r \lambda^r, \tag{2.2}$$

where f_r is the derivative $\partial f/\partial x^r$. The second derivative, $\partial^2 f/\partial s^2$, can be obtained by differentiating the right hand side of (2.2) *covariantly* with respect to x^t and the form (1.1), multiplying by λ^t, and summing on t. We obtain

$$\frac{\partial^2 f}{\partial s^2} = f_{rt}\lambda^r\lambda^t + f_r\lambda^r{}_{,t}\lambda^t \tag{2.3}$$

where f_{rt} and $\lambda^r{}_{,t}$ are the covariant derivatives of f_r and λ^r respectively.

If C is a geodesic, and if we denote its arc by u^0, we must have†

$$\frac{d^2x^r}{(du^0)^2} + \left\{ \begin{matrix} r \\ ij \end{matrix} \right\} \frac{dx^i}{du^0} \frac{dx^j}{du^0} = 0 \quad (r = 0,1,2, \cdots, n). \tag{2.4}$$

This may be written in the equivalent form‡

$$\lambda^r{}_{,i}\lambda^i = 0, \tag{2.5}$$

where $\lambda^r{}_{,i}$ is the covariant derivative of λ^r with respect to x^i and the form (1.1). In equations (2.3) the second term drops out by virtue of equations (2.5), so that we have

$$\frac{\partial^2 f}{(\partial u^0)^2} = f_{rs}\lambda^r\lambda^s. \tag{2.6}$$

By induction it follows that

$$\frac{\partial^m f}{(\partial u^0)^m} = f_{r_1 r_2 \cdots r_m}\lambda^{r_1}\lambda^{r_2} \cdots \lambda^{r_m} \qquad (m = 1,2, \cdots). \tag{2.7}$$

---

* If C is not a *null* curve, that is, if, for C, $a_{ij}dx^i dx^j \neq 0$, the arc is defined, $ds^2 = |a_{ij}dx^i dx^j|$. If C is null, s must be interpreted as a parameter along the curve. Cf. Eisenhart, loc. cit.

† Bianchi, 2d edition, loc. cit., p. 334, or any of the other works on this subject.

‡ Cf. Levi-Civita, loc. cit., p. 291, or any of the similar works.

Hence we have the following theorem:

The mth directional derivative of any function f along a geodesic is obtained by taking the inner product of the mth covariant derivative of f with $\lambda^{r_1}\lambda^{r_2}\cdots\lambda^{r_m}$, *where* λ^r *is the vector tangent to the geodesic.*

In particular, it follows that, if $-x^i$ is taken for the function f we must have

$$-\frac{d^m x^i}{(du^0)^m} = \Gamma^i_{r_1\ldots r_m}\frac{dx^{r_1}}{du^0}\cdots\frac{dx^{r_m}}{du^0} \quad (i = 0,1,2,\cdots,n), \tag{2.8}$$

where the functions $\Gamma^i_{r_1\ldots r_m}$ are formed from $-x^i$ by the formal process of covariant differentiation,

$$\begin{aligned}
\Gamma^i_{r_1} &= \left.\begin{matrix} 0, & r_1 = i \\ -1, & r_1 \neq i \end{matrix}\right\}, \\
\Gamma^i_{rs} &= \frac{\partial\Gamma^i_r}{\partial x^s} - \Gamma^i_j\left\{\begin{matrix} j \\ r\,s \end{matrix}\right\} \equiv \left\{\begin{matrix} i \\ r\,s \end{matrix}\right\}, \\
\Gamma^i_{rst} &= \frac{\delta\Gamma^i_{rs}}{\delta x^t} - \Gamma^i_{js}\left\{\begin{matrix} j \\ r\,t \end{matrix}\right\} - \Gamma^i_{rj}\left\{\begin{matrix} j \\ s\,t \end{matrix}\right\}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{2.9}$$

Hence that solution of equations (2.4) which has initial values y^i and ξ^i,

$$x^i\Big|_{u^0=0} = y^i, \qquad \frac{dx^i}{du^0}\Big|_{u^0=0} = \xi^i,$$

must be*

$$x^i = y^i + \xi^i u^0 - \frac{1}{2!}\left\{\begin{matrix} i \\ r\,s \end{matrix}\right\}\xi^r\xi^s(u^0)^2 - \frac{1}{3!}\Gamma^i_{rst}\xi^r\xi^s\xi^t(u_0)^3 - \cdots, \tag{2.10}$$

where $\left\{\begin{smallmatrix} i \\ r\,s \end{smallmatrix}\right\}$ and the Γ's are evaluated for the point y^i.

3. We consider an arbitrary tensor $T_{ij\ldots k}$ expanded in a power series, for simplicity, in one variable,

$$T_{ij\ldots k} = \overset{(0)}{T}_{ij\ldots k} + \overset{(1)}{T}_{ij\ldots k}(x^0) + \overset{(2)}{T}_{ij\ldots k}(x^0)^2 + \cdots \quad (i,\ j,\ k = 0,\ 1,\ 2,\ \cdots,\ n). \tag{3.1}$$

If we make the transformation of coördinates defined by the equations

$$\begin{aligned} x^0 &= \bar{x}^0, \\ x^i &= x^i(\bar{x}^1,\ \bar{x}^2,\ \cdots,\ \bar{x}^n) \qquad (i = 1,\ 2,\ \cdots,\ n), \end{aligned} \tag{3.2}$$

* Cf. Veblen, loc. cit., p. 192.

the components $\bar{T}_{ij\cdots k}$ in the new coördinate system undergo the usual tensor transformations, which in this case, and for a covariant tensor, reduce to

$$\bar{T}_{\alpha\beta\cdots\gamma} = \sum_{\lambda,\mu,\nu=1}^{n} T_{\lambda\mu\cdots\nu} \frac{\partial x^\lambda}{\partial \bar{x}^\alpha} \frac{\partial x^\mu}{\partial \bar{x}^\beta} \cdots \frac{\partial x^\nu}{\partial \bar{x}^\gamma} \tag{3.3}$$

$$(\alpha, \beta, \gamma = 1, 2, \cdots, n).$$

Since x^0 appears only in T and not in the derivatives, it follows by differentiation of (3.3) with respect to x^0 that

$$\bar{T}^{(m)}_{\alpha\beta\cdots\gamma} = \sum_{\lambda,\mu,\nu}^{1\cdots n} T^{(m)}_{\lambda\mu\cdots\nu} \frac{\partial x^\lambda}{\partial \bar{x}^\alpha} \frac{\partial x^\mu}{\partial \bar{x}^\beta} \cdots \frac{\partial x^\nu}{\partial \bar{x}^\gamma}, \tag{3.4}$$

that is, that $T^{(m)}_{\lambda\mu\cdots\nu}$ are the components of a covariant tensor in $x^0=0$. The similar result for a mixed or contravariant tensor or invariant obviously is valid. If the expansion (3.1) is in a series of several variables, the result holds for the variety of the remaining variables. Putting this in the form of a theorem, we have

If $T_{ij\cdots k}$ and $\bar{T}_{ij\cdots k}$ are the components of a tensor in the coördinate systems x and $\bar{x}$ respectively, and if these coördinates are related by equations

$$x^\alpha = \bar{x}^\alpha \qquad (\alpha = 0,1,2, \cdots, p-1),$$

$$x^\mu = x^\mu(\bar{x}^p, \bar{x}^{p+1}, \cdots, \bar{x}^n) \qquad (\mu = p, p+1, \cdots, n),$$

then the coefficients in the expansion of $T_{\lambda\mu\cdots\nu}$ in a power series in $x^0, x^1, \cdots, x^{p-1}$ are the components of a tensor in the subspace $x^\alpha=$constant $(\alpha=0, 1, \cdots, p-1)$, and the components of this tensor in the coördinates $\bar{x}^p, \cdots, \bar{x}^n$ are the coefficients in the expansion of $\bar{T}_{\lambda\mu\cdots\nu}$ in a power series in $x^0, x^1, \cdots, x^{p-1}$.

Part II

4. We consider a hypersurface of n dimensions denoted by V_n immersed in V_{n+1} with the quadratic form (1.1). Such a spread is defined by the equations

$$x^i = f^i(u^1, u^2, \cdots, u^n) \qquad (i = 0,1,2, \cdots, n), \tag{4.1}$$

where the jacobian matrix

$$J = \left\| \frac{\partial f^i}{\partial u^\alpha} \right\| \qquad \begin{pmatrix} i = 0,1,2, \cdots, n \\ \alpha = 1,2, \cdots, n \end{pmatrix} \tag{4.2}$$

is of rank n. The fundamental form of V_n is given by

$$\varphi' = g_{\alpha\beta} du^\alpha du^\beta \tag{4.3}$$

where

$$g_{\alpha\beta} = a_{ij}\frac{\partial f^i}{\partial u^\alpha}\frac{\partial f^j}{\partial u^\beta} \qquad (\alpha,\beta = 1,2,\cdots,n). \tag{4.4}$$

In equations (4.3), α and β are summed from 1 to n, and throughout the remainder of this paper Greek letters will be reserved for such summations, while Latin letters as i and j in (4.4) indicate a summation from 0 to n.

When the form (1.1) is positive definite, (4.3) will be likewise; but when no assumption is made concerning the definiteness of (1.1), it is necessary to assume explicitly that the functions f^i are such that the discriminant g of the form (4.3),

$$g = |g_{\alpha\beta}|, \tag{4.5}$$

is different from zero.

As usual, $g^{\alpha\beta}$ denotes the cofactor of $g_{\alpha\beta}$ in g divided by g itself, so that

$$g_{\alpha\beta}g^{\beta\gamma} = \delta_\alpha^\gamma \qquad (\alpha,\gamma = 1,2,\cdots,n), \tag{4.6}$$

where δ_α^γ is defined by equations (1.4).

Eisenhart* has shown that the condition $g\neq 0$ and the condition that the vector normal to V_n be not null are equivalent; hence if we denote the components of this vector by ξ^i they may be chosen so that

$$a_{ij}\xi^i\xi^j = e, \tag{4.7}$$

where e is plus or minus one.

The investigations we wish to make will be facilitated if instead of the general coördinate system $x^0, x^1, \cdots, x^n$ we make use of a coördinate system $u^0, u^1, \cdots, u^n$, where the u^0 coördinate of a point measures the distance of that point from V_n along a geodesic normal to V_n and the coördinates $u^1, u^2, \cdots, u^n$ are any coördinates in V_n. In these coördinates we find that†

$$c_{11} = e, \qquad c_{0\alpha} = c_{\alpha 0} = 0 \qquad (\alpha = 1,2,\cdots,n), \tag{4.8}$$

where e is defined by equation (4.7), and hence the form (1.1) becomes

$$\varphi = e(du^0)^2 + c_{\alpha\beta}du^\alpha du^\beta. \tag{4.9}$$

---

* Eisenhart, loc. cit., p. 144.

† Cf. Bianchi, 2d edition, loc. cit., p. 336, when the form (1.1) is assumed positive definite. Eisenhart, loc. cit., p. 146, treats the general case.

The equation of V_n in these coördinates is $u^0=0$, and equations (4.4) become

$$g_{\alpha\beta} = c_{\alpha\beta}(0,u^1, \cdots, u^n) \qquad (\alpha,\beta = 1,2, \cdots,n). \tag{4.10}$$

If c^{ij} is the cofactor of c_{ij} in the determinant $c=|c_{ij}|$ divided by c itself,

$$g^{\alpha\beta} = c^{\alpha\beta}(0,u^1, \cdots,u^n) \qquad (\alpha,\beta = 1,2, \cdots, n). \tag{4.11}$$

We observe further that the components in the coördinates u^i of the vector normal to V_n are given by

$$\xi^0 = 1, \qquad \xi^\alpha = 0 \qquad (\alpha = 1,2, \cdots,n). \tag{4.12}$$

5. Bianchi considers the variation in the fundamental form of V_{n+1} as one passes from V_n to a hypersurface geodesically parallel to V_n. Letting ϵ be the constant distance between these two hypersurfaces, he finds that*

$$\delta\varphi = \epsilon\left(\frac{\partial c_{\alpha\beta}}{\partial u^0}\right)_0 du^\alpha du^\beta, \tag{5.1}$$

and he defines functions $\Omega_{\alpha\beta}$ by the equations

$$\Omega_{\alpha\beta} = -\frac{1}{2}\left.\frac{\partial c_{\alpha\beta}}{\partial u^0}\right|_{u^0=0}. \tag{5.2}$$

The two quadratic forms that we are thus led to, $g_{\alpha\beta}du^\alpha du^\beta$ and $\Omega_{\alpha\beta}du^\alpha du^\beta$, known respectively as the *first* and *second fundamental forms* of the hypersurface, are of prime importance in the study of the variety. But their coefficients $g_{\alpha\beta}$ and $\Omega_{\alpha\beta}$ are only the first two terms in the expansion of $c_{\alpha\beta}$ in a power series in u^0,

$$c_{\alpha\beta} = g_{\alpha\beta} - 2\Omega_{\alpha\beta}u^0 + \overset{(2)}{c}_{\alpha\beta}(u^0)^2 + \overset{(3)}{c}_{\alpha\beta}(u^0)^3 + \cdots, \tag{5.3}$$

so it is natural to inquire into the nature of the remaining terms.

That the functions $\overset{(m)}{c}_{\alpha\beta}$ are of an intrinsic nature is shown by the following theorem:

The functions $\overset{(m)}{c}_{\alpha\beta}$ are tensors in the hypersurface and invariants in the space.

The proof of the first statement is a direct consequence of the theorem of § 3. For the second part we observe that if we had, in V_{n+1}, two general coördinate systems x and $\bar{x}$ it would follow that

$$a_{ij}\frac{\partial x^i}{\partial u^\alpha}\frac{\partial x^j}{\partial u^\beta} = \left(\bar{a}_{rs}\frac{\partial \bar{x}^r}{\partial x^i}\frac{\partial \bar{x}^s}{\partial x^j}\right)\frac{\partial x^i}{\partial u^\alpha}\frac{\partial x^j}{\partial u^\beta} = \bar{a}_{rs}\frac{\partial \bar{x}^r}{\partial u^\alpha}\frac{\partial \bar{x}^s}{\partial u^\beta},$$

* Bianchi, 2d edition, loc. cit., p. 360.

and hence

$$c_{\alpha\beta} = a_{ij}\frac{\partial x^i}{\partial u^\alpha}\ \frac{\partial x^j}{\partial u^\beta} \qquad (\alpha,\beta = 1, 2, \cdots, n)$$

is an invariant in V_{n+1}, and consequently the functions $\overset{(m)}{c}_{\alpha\beta}$ are invariants in V_{n+1}.

6. We devote the remainder of this paper to the determination of these tensors and to some applications to particular spaces and hypersurfaces.

From the definitions of the Christoffel symbols in equations (1.5) and (1.6), we have that, for the fundamental form (4.9),

$$(6.1)\quad [00,0] = [00,\alpha] = [0\alpha,0] = 0, \quad [0\alpha,\beta] = \frac{1}{2}\frac{\partial c_{\alpha\beta}}{\partial u^0} \quad (\alpha,\beta = 1,2, \cdots,n)$$

and hence that

$$(6.2)\quad \begin{Bmatrix} 0 \\ 0\ 0 \end{Bmatrix} = \begin{Bmatrix} \alpha \\ 0\ 0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0\ \alpha \end{Bmatrix} = 0, \begin{Bmatrix} \alpha \\ 0\ \beta \end{Bmatrix} = \frac{1}{2}\, c^{\alpha\gamma}\frac{\partial c_{\beta\gamma}}{\partial u^0} \qquad (\alpha,\beta = 1,2, \cdots,n).$$

If we evaluate $R_{0\alpha\beta 0}$ from equations (1.7), we are able to solve for the second derivative, obtaining that

$$(6.3)\qquad \frac{\partial^2 c_{\alpha\beta}}{(\partial u^0)^2} = 2R_{0\alpha\beta 0} + \frac{1}{2}\, c^{\gamma\delta}\frac{\partial c_{\alpha\gamma}}{\partial u^0}\ \frac{\partial c_{\beta\delta}}{\partial u^0} \qquad (\alpha,\beta = 1,2, \cdots,n),$$

and hence, by virtue of (4.11) and (5.2),

$$(6.4)\qquad \overset{(2)}{c}_{\alpha\beta} = R_{0\alpha\beta 0} + \Omega^\gamma_\alpha \Omega_{\beta\gamma},$$

where

$$(6.5)\qquad \Omega^\gamma_\alpha = g^{\gamma\beta}\Omega_{\alpha\beta} = g^{\gamma\beta}\Omega_{\beta\alpha},$$

and it is understood in (6.4) that $R_{0\alpha\beta 0}$ is evaluated on the hypersurface. For a general coördinate system, $x^0, x^1, \cdots, x^n$, in the space, the third fundamental tensor is given by

$$(6.6)\qquad \overset{(2)}{c}_{\alpha\beta} = \overline{R}_{ijkl}\frac{\partial x^j}{\partial u^\alpha}\ \frac{\partial x^k}{\partial u^\beta}\xi^i\xi^l + \Omega^\gamma_\alpha\Omega_{\beta\gamma},$$

where the $\overline{R}$'s here are the Riemann tensors with respect to the fundamental form of the x's.

We can obtain $\partial^3 c_{\alpha\beta}/(\partial u^0)^3$ by differentiating equations (6.3),

$$(6.7)\qquad \frac{\partial^3 c_{\alpha\beta}}{(\partial u^0)^3} = 2\frac{\partial}{\partial u^0}R_{0\alpha\beta 0} + \frac{1}{2}\frac{\partial}{\partial u^0}\left(c^{\gamma\delta}\ \frac{\partial c_{\alpha\gamma}}{\partial u^0}\ \frac{\partial c_{\beta\delta}}{\partial u^0}\right).$$

The ordinary derivative of $R_{0\alpha\beta 0}$ has not an invariantive significance, and so it is preferable to replace it by the covariant derivative $R_{0\alpha\beta 0,0}$. From the definition, simplified by virtue of (6.2), we obtain that

$$\frac{\partial}{\partial u^0} R_{0\alpha\beta 0} = R_{0\alpha\beta 0,0} + \frac{1}{2} c^{\mu\nu} \frac{\partial c_{\alpha\nu}}{\partial u^0} R_{0\mu\beta 0} + \frac{1}{2} c^{\mu\nu} \frac{\partial c_{\beta\nu}}{\partial u^0} R_{0\alpha\nu 0}. \tag{6.8}$$

Furthermore, from

$$c^{\alpha i} c_{i\beta} = c^{\alpha\gamma} c_{\gamma\beta} = \delta^{\alpha}_{\beta} \qquad (\alpha, \beta = 1, 2, \cdots, n) \tag{6.9}$$

it follows that

$$c^{\alpha\gamma} \frac{\partial c_{\gamma\beta}}{\partial u^0} = - c_{\gamma\beta} \frac{\partial c^{\alpha\gamma}}{\partial u^0} . \tag{6.10}$$

Hence equations (6.3) are equivalent to

$$c^{\delta\gamma} \frac{\partial^2 c_{\alpha\gamma}}{(\partial u^0)^2} = 2R^{\delta}_{0\alpha 0} - \frac{1}{2} \frac{\partial c_{\alpha\gamma}}{\partial u^0} \frac{\partial c^{\delta\gamma}}{\partial u^0},$$

from which we obtain that

$$\frac{\partial}{\partial u^0}\left(c^{\gamma\delta} \frac{\partial c_{\alpha\gamma}}{\partial u^0} \right) = 2R^{\delta}_{0\alpha 0} + \frac{1}{2} \frac{\partial c_{\alpha\gamma}}{\partial u^0} \frac{\partial c^{\delta\gamma}}{\partial u^0} . \tag{6.11}$$

Substituting (6.8), (6.11), and (6.3) in (6.7) and making a slight simplification due to the equalities

$$R_{0\alpha\beta 0} = R_{0\beta\alpha 0}, \quad R^{\beta}_{0\alpha 0} = c^{\beta\gamma} R_{0\gamma\alpha 0} = c^{\beta\gamma} R_{0\alpha\gamma 0}, \tag{6.12}$$

we obtain that

$$\frac{\partial^3 c_{\alpha\beta}}{(\partial u^0)^3} = 2R_{0\alpha\beta 0,0} + 2c^{\mu\nu} \left(\frac{\partial c_{\mu\alpha}}{\partial u^0} R_{0\nu\beta 0} + \frac{\partial c_{\mu\beta}}{\partial u^0} R_{0\alpha\nu 0} \right). \tag{6.13}$$

To obtain the fourth derivative in intrinsic form, we observe that because of equation (6.2) we are able to write

$$\frac{\partial R_{0\alpha\beta 0,0}}{\partial u^0} = R_{0\alpha\beta 0,00} + \frac{1}{2} c^{\mu\nu} \left\{ \frac{\partial c_{\nu\alpha}}{\partial u^0} R_{0\mu\beta 0,0} + \frac{\partial c_{\nu\beta}}{\partial u^0} R_{0\alpha\mu 0,0} \right\}. \tag{6.14}$$

Furthermore, in the consideration of the term

$$c^{\mu\nu} \frac{\partial c_{\mu\alpha}}{\partial u^0} R_{0\nu\beta 0}$$

of equations (6.13), differentiation of the first two members by application of equation (6.11) introduces, among others, the term

$$\frac{\partial c_{\mu\alpha}}{\partial u^0}\ \frac{\partial c^{\mu\nu}}{\partial u^0}\, R_{0\nu\beta 0},$$

while the application of formula (6.8) to the differentiation of $R_{0\nu\beta 0}$ introduces

$$c^{\mu\nu}\frac{\partial c_{\mu\alpha}}{\partial u^0}\, c^{\sigma\tau}\,\frac{\partial c_{\tau\nu}}{\partial u^0}\, R_{0\sigma\beta 0}\ ;$$

because of equations (6.10) these expressions are equal numerically, but opposite in sign; the two similar terms obtained in the differentiation of

$$c^{\mu\nu}\frac{\partial c_{\nu\beta}}{\partial u^0}\, R_{0\alpha\nu 0}$$

likewise cancel each other. These are the only cancellations occurring, so that we have

$$
\begin{aligned}
\frac{\partial^4 c_{\alpha\beta}}{(\partial u^0)^4} &= 2R_{0\alpha\beta 0,00} + 8R^{\mu}_{0\alpha 0}\, R_{0\mu\beta 0} \\
&\quad + 3c^{\nu\mu}\left(\frac{\partial c_{\mu\alpha}}{\partial u^0}\, R_{0\nu\beta 0,0} + \frac{\partial c_{\mu\beta}}{\partial u^0}\, R_{0\alpha\nu 0,0}\right) \qquad (6.15)\\
&\quad + 2R_{0\lambda\mu 0}c^{\lambda\sigma}\,\frac{\partial c_{\sigma\alpha}}{\partial u^0}\, c^{\mu\tau}\,\frac{\partial c_{\tau\beta}}{\partial u^0}\,.
\end{aligned}
$$

Similarly, we obtain $\partial^5 c_{\alpha\beta}/(\partial u^0)^5$ by differentiation of (6.15), but instead, let us turn to the general derivative to evaluate it.

7. For the consideration of the general coefficient it will be desirable to introduce a slightly more convenient notation, in part necessary for the sake of clearness. We can see that the covariant derivatives of $R_{0\alpha\beta 0}$ with respect to u^0 and of all orders will appear in the following pages and we shall denote by $R_{0\alpha\beta 0,[l]}$ and $R^{\alpha}_{0\beta 0,[l]}$ the lth covariant derivative with respect to u^0 of $R_{0\alpha\beta 0}$ and $R^{\alpha}_{0\beta 0}$ respectively. Thus,

$$R_{0\alpha\beta 0,[3]} \equiv R_{0\alpha\beta 0,000}, \qquad R^{\alpha}_{0\beta 0,[4]} \equiv R^{\alpha}_{0\beta 0,0000}.$$

For symmetry in our formulas, we shall interpret $R_{0\alpha\beta 0,[0]}$ as merely $R_{0\alpha\beta 0}$, that is,

$$R_{0\alpha\beta 0,[0]} \equiv R_{0\alpha\beta 0}, \qquad R^{\alpha}_{0\beta 0,[0]} \equiv R^{\alpha}_{0\beta 0}.$$

With this notation we are able to write the general formula suggested by equations (6.8) and (6.14), namely

$$\frac{\partial}{\partial u^0} R_{0\alpha\beta 0,[l]} = R_{0\alpha\beta 0,[l+1]} + \frac{1}{2} c^{\mu\nu} \left(\frac{\partial c_{\nu\alpha}}{\partial u^0} R_{0\mu\beta 0,[l]} + \frac{\partial c_{\nu\beta}}{\partial u^0} R_{0\alpha\mu 0,[l]} \right). \tag{7.1}$$

This follows immediately from the definition of covariant differentiation when the Christoffel symbols are subject to equations (6.2). Similarly we have for the covariant derivative of the four-index symbol of Riemann of the second kind, that

$$\frac{\partial R^{\alpha}_{0\beta 0,[l]}}{\partial u^0} = R^{\alpha}_{0\beta 0,[l+1]} + \frac{1}{2} c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} R^{\alpha}_{0\mu 0,[l]} - \frac{1}{2} c^{\mu\alpha} \frac{\partial c_{\mu\nu}}{\partial u^0} R^{\nu}_{0\beta 0,[l]}. \tag{7.2}$$

In the following discussion there will arise contracted products of several R's. We introduce the notation $\mathcal{R}_{[k_1 k_2 \cdots k_N]\alpha\beta}$, defining it by the equation

$$\mathcal{R}_{[k_1 \cdots k_N]\alpha\beta} \equiv R^{\mu_1}_{0\mu_2 0,[k_1]} R^{\mu_2}_{0\mu_3 0,[k_2]} \cdots R^{\mu_{N-2}}_{0\mu_{N-1} 0,[k_{N-2}]} R^{\mu_{N-1}}_{0\alpha 0,[k_{N-1}]} R_{0\mu_1\beta 0,[k_N]}. \tag{7.3}$$

We note that N may take on any value 2, 3, $\cdots$, while the k's may take on 0, 1, 2, $\cdots$. For further clarity let us observe that

$$\mathcal{R}_{[01]\alpha\beta} = R^{\mu}_{0\alpha 0} R_{0\mu\beta 0,0},$$

$$\mathcal{R}_{[210]\alpha\beta} = R^{\mu}_{0\nu 0,00} R^{\nu}_{0\alpha 0,0} R_{0\mu\beta 0}.$$

From the definition, from equations (6.12) and from the equations derived from (6.12) by covariant differentiation, it follows that

$$\mathcal{R}_{[k_1 k_2 \ldots k_N]\alpha\beta} \equiv \mathcal{R}_{[k_{N-2} k_{N-3} \cdots k_1 k_N k_{N-1}]\beta\alpha}. \tag{7.4}$$

If we differentiate equations (7.3) covariantly with respect to u^0, we find that

$$\text{Covariant Deriv. of } \mathcal{R}_{[k_1 \cdots k_N]\alpha\beta} = \sum_{i=1}^{N} \mathcal{R}_{[k_1 \cdots k_{i-1} k_i+1 k_{i+1} \cdots k_N]\alpha\beta}. \tag{7.5}$$

Again, if we evaluate $(\partial/\partial u^0)\mathcal{R}_{[k_1 \cdots k_N]\alpha\beta}$ by means of equations (7.1) and (7.2) we find that the term $R^{\mu_i}_{0\mu_{i+1}0,[k_i]}$ of $\mathcal{R}$ contributes the single term $R^{\mu_i}_{0\mu_{i+1}0,[k_i+1]}$, for the other two terms which arise from the derivative of $R^{\mu_i}_{0\mu_{i+1}0,[k_i]}$ are cancelled by two similar terms one of which arises from the differentiation of $R^{\mu_{i-1}}_{0\mu_i 0,[k_{i-1}]}$ and the other from the differentiation of $R^{\mu_{i+1}}_{0\mu_{i+2}0,[k_{i+1}]}$; hence we have that

$$\begin{aligned} \frac{\partial}{\partial u^0} \mathcal{R}_{[k_1 \ldots k_N]\alpha\beta} = {} & \sum_{i=1}^{n} \mathcal{R}_{[k_1 \cdots k_{i-1} k_i+1 k_{i+1} \cdots k_N]\alpha\beta} \\ & + \frac{1}{2} c^{\mu\nu} \frac{\partial c_{\nu\alpha}}{\partial u^0} \mathcal{R}_{[k_1 \cdots k_N]\alpha\mu} + \frac{1}{2} c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathcal{R}_{[k_1 \cdots k_N]\alpha\mu}. \end{aligned} \tag{7.6}$$

We thus see that the differentiation of $\mathfrak{R}_{[k_1 \dots k_N]\alpha\beta}$ introduces terms of the type

$$c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\alpha\mu}$$

and accordingly we should like to have the expression for the derivative of a term of this type. Equations (6.11) and (7.6) enable us to obtain it; here again there will be a cancellation, the second term of (6.11) annulling one of the terms introduced by the differentiation of $\mathfrak{R}$. Observing that $R^{\mu}_{0\beta 0} \mathfrak{R}_{[k^1 \dots k_N]\alpha\mu} = \mathfrak{R}_{[k_N k_1 \dots k_{N-1} 0]\alpha\beta}$, we are able to write

$$\frac{\partial}{\partial u^0}\left\{ c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\alpha\beta} \right\} = 2\mathfrak{R}_{[k_N k_1 \dots k_{N-1} 0]\alpha\beta} + c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \sum_{i=1}^{N} \mathfrak{R}_{[k_1 \dots k_{i-1} k_i + 1 k_{i+1} \dots k_N]\alpha\beta} + \frac{1}{2} c^{\mu\nu} c^{\sigma\tau} \frac{\partial c_{\nu\beta}}{\partial u^0} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\tau\mu}. \tag{7.7}$$

Either from equations (7.4) or directly, we have the similar result that

$$\frac{\partial}{\partial u^0}\left\{ c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\mu\alpha} \right\} = 2\mathfrak{R}_{[k_1 \dots k_{N-1} 0 k_N]\beta\alpha} + c^{\nu\mu} \frac{\partial c_{\nu\beta}}{\partial u^0} \sum_{i=1}^{N} \mathfrak{R}_{[k_1 \dots k_{i-1} k_i + 1 k_{i+1} \dots k_N]\mu\alpha} + \frac{1}{2} c^{\mu\nu} c^{\lambda\sigma} \frac{\partial c_{\nu\beta}}{\partial u^0} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\mu\lambda}. \tag{7.8}$$

Finally, we observe that we need a formula for the derivative of

$$c^{\lambda\sigma} c^{\mu\nu} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\lambda\mu}.$$

This we obtain from a combination of equations (6.11) with either (7.7) or (7.8). Here again the term arising from the second term of (6.11) will cancel the one arising from the last term of equations (7.7) (or (7.8)). The result therefore is

$$\frac{\partial}{\partial u^0}\left\{ c^{\lambda\sigma} c^{\mu\nu} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_N]\lambda\mu} \right\} = 2c^{\mu\nu} \frac{\partial c_{\nu\beta}}{\partial u^0} \mathfrak{R}_{[k_1 \dots k_{N-1} 0 k_N]\alpha\mu} + 2c^{\lambda\sigma} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \mathfrak{R}_{[k_N k_1 \dots k_{N-1} 0]\lambda\beta} + c^{\lambda\sigma} c^{\mu\nu} \frac{\partial c_{\sigma\alpha}}{\partial u^0} \frac{\partial c_{\nu\beta}}{\partial u^0} \sum_{i=1}^{N} \mathfrak{R}_{[k_1 \dots k_{i-1} k_i + 1 k_{i+1} \dots k_N]\lambda\mu}. \tag{7.9}$$

The formulas here obtained, namely (7.6), (7.7), (7.8) and (7.9), indicate the general character of $\partial^m c_{\alpha\beta}/(\partial u^0)^m$ and lead us to write the following expression:

$$
\begin{aligned}
\frac{\partial^m c_{\alpha\beta}}{(\partial u^0)^m} = {} & 2R_{0\alpha\beta 0,[m-2]} + \sum_{\sigma,N}^{m} \overset{(m)}{D}_{\sigma_1\sigma_2\ldots\sigma_N} \mathcal{R}_{[\sigma_1\ldots\sigma_N]\alpha\beta} \\
& + c^{\lambda\mu}\frac{\partial c_{\lambda\alpha}}{\partial u^0}\left\{(m-1)R_{0\mu\beta 0,[m-3]} + \sum_{\tau,N}^{m-1} \overset{(m)}{E}_{\tau_1\ldots\tau_N}\mathcal{R}_{[\tau_1\ldots\tau_N]\mu\beta}\right\} \\
& + c^{\lambda\mu}\frac{\partial c_{\lambda\alpha}}{\partial u^0}\left\{(m-1)R_{0\mu\alpha 0,[m-3]} + \sum_{\tau,N}^{m-1} \overset{(m)}{\overline{E}}_{\tau_1\ldots\tau_N}\mathcal{R}_{[\tau_1\ldots\tau_N]\mu\alpha}\right\} \\
& + c^{\nu\mu}c^{\nu\sigma}\frac{\partial c_{\lambda\alpha}}{\partial u^0}\frac{\partial c_{\sigma\beta}}{\partial u^0}\left\{\frac{m(m-3)}{2}R_{0\mu\nu 0,[m-4]} + \sum_{\pi,N}^{m-2} \overset{(m)}{F}_{\pi_1\ldots\pi_N}\mathcal{R}_{[\pi_1\ldots\pi_N]\mu\nu}\right\}
\end{aligned}
\tag{7.10}
$$

where the D's, E's, and F's are constants to be determined, and where $\sum_{\sigma,N}^{m}$ represents the sum over $N=2, 3, 4 \cdots$ and $\sigma_i=0, 1, 2, \cdots$ and such that $2N+\sigma_1+\sigma_2+\cdots+\sigma_N=m$.

Since $\partial^m c_{\alpha\beta}/(\partial u^0)^m$ is symmetric in α and β, the right hand side of (7.10) must be also, and independently of the R's; this gives us conditions which by virtue of equation (7.4) become

$$
\begin{aligned}
\overset{(m)}{D}_{\sigma_1\ldots\sigma_N} &= \overset{(m)}{D}_{\sigma_{N-2}\sigma_{N-3}\ldots\sigma_1\sigma_N\sigma_{N-1}}, \\
\overset{(m)}{\overline{E}}_{\tau_1\ldots\tau_N} &= \overset{(m)}{E}_{\tau_1\ldots\tau_N}, \\
\overset{(m)}{F}_{\pi_1\ldots\pi_N} &= F_{\pi_{N-2}\ldots\pi_1\pi_2\pi_{N-1}}.
\end{aligned}
\tag{7.11}
$$

If we form $\partial^{m+1}c_{\alpha\beta}/(\partial u^0)^{m+1}$ by differentiating (7.10) with respect to u^0 making use of the four fundamental formulas (7.6), (7.7), (7.8) and (7.9) and of equations (7.4) and (7.11), and compare the result with the equation obtained from (7.10) by writing $m+1$ for m, we obtain the justification for the precise numerical coefficients of (7.10) and at the same time we find recursion formulas for the D's, E's, and F's. To facilitate this differentiation and the collection of the results, we observe that there are three types of terms in (7.10), those in which $\partial c_{\mu\nu}/\partial u^0$ does not appear, those linear in these derivatives, and finally, those quadratic in them. Those of the first type can arise only from those of the first or second type, those of the second type arise from all three types, but those of the third type arise only from those of the second and third types. With these observations the recursion formulas follow immediately from equations (7.6) to (7.9); they are

$$D^{m+1}_{\rho_1\cdots\rho_N} = \sum_{i=1}^{N} D^{(m)}_{\rho_1\cdots\rho_{i-1}\,\rho_{i-1}\rho_{i+1}\cdots\rho_N} + 2(m-1)\delta^{m-3}_{\rho_N}\,\delta^{0}_{\rho_{N-1}}$$

$$+\,2E^{(m)}_{\rho_1\cdots\rho_{N-2}\rho_N}\delta^{0}_{\rho_{N-1}} + 2(m-1)\delta^{m-3}_{\rho_{N-1}}\,\delta^{0}_{\rho_N} \tag{7.12}$$

$$+\,2E^{(m)}_{\rho_{N-2}\rho_{N-3}\cdots\rho_1\rho_N}\delta^{0}_{\rho_N},$$

$$E^{(m+1)}_{\sigma_1\cdots\sigma_N} = \tfrac{1}{2}\,D^{(m)}_{\sigma_1\cdots\sigma_N} + \sum_{i=1}^{N} E^{(m)}_{\sigma_1\ldots\sigma_i-1\ldots\sigma_N}$$

$$+m(m-3)\delta^{m-4}_{\sigma_{N-1}}\delta^{0}_{\sigma_N} + 2F_{\sigma_1\ldots\sigma_{N-1}\sigma_1}\delta^{0}_{\sigma_N}, \tag{7.13}$$

$$F^{(m+1)}_{\tau_1\ldots\tau_N} = \sum_{i=1}^{N} F^{(m)}_{\tau_1\ldots\tau_i-1\ldots\tau_N} + \tfrac{1}{2}\,E^{(m)}_{\tau_1\ldots\tau_N}$$

$$+\,\tfrac{1}{2}\,E^{(m)}_{\tau_{N-2}\tau_{N-3}\ldots\tau_1\tau_N\tau_{N-1}}, \tag{7.14}$$

where the δ's are defined by equations (1.4) and where a D, E, or F with a negative subscript is to be interpreted as zero.

Summing up these results in the form of a theorem, we have the following:

The mth derivative of $c_{\alpha\beta}$ with respect to u^0 is expressible as a non-homogeneous polynomial of the second degree in the first derivatives $\partial c_{\lambda\mu}/\partial u^0$ ($\lambda, \mu = 1, 2, \cdots, n$) whose coefficients are polynomials in the Riemann tensor and its covariant derivatives of orders less than and equal to $m-2$.

8. The fundamental tensors in questions, $\overset{(m)}{c}_{\alpha\beta}$, are the derivatives evaluated on the hypersurface $u^0=0$. We recall the definition of $\Omega_{\alpha\beta}$ and Ω^{α}_{β} from equations (5.2) and (6.5) respectively, and the relations (4.10) and (4.11) between $g_{\alpha\beta}$ and $c_{\alpha\beta}$. We observe further that if $\overline{R}_{ijkl,r\ldots s}$ is the R in a general coördinate system, we must have

$$R_{0\alpha\beta0,[l]} = \overline{R}_{hijk,r_1\ldots r_l}\frac{\partial x^i}{\partial u^\alpha}\frac{\partial x^j}{\partial u^\beta}\xi^h\xi^k\xi^{r_1}\cdots\xi^{r_l} \tag{8.1}$$

where $\xi^i = \partial x^i/\partial u^0$ is the normal vector to V_n. Hence this theorem follows:

The fundamental tensors of a hypersurface are expressible in terms of the first two fundamental tensors of the hypersurface and of the Riemann tensor of the space and its covariant derivatives, evaluated on the hypersurface.

For reference we shall give the explicit values of the first five of these tensors, the first three being obtained directly from equations (6.4), (6.13) and (6.15) and the others by repeated application of the recursion formulas

(7.12), (7.13) and (7.14)*:

$$\overset{(2)}{c}_{\alpha\beta} = R_{0\alpha\beta0} + \Omega^{\gamma}_{\alpha}\Omega_{\gamma\beta},$$

$$\overset{(3)}{c}_{\alpha\beta} = \frac{2}{3\,!} R_{0\alpha\beta0,0} - \frac{4}{3\,!}(R_{0\alpha\gamma0}\Omega^{\gamma}_{\beta} + R_{0\gamma\beta0}\Omega^{\gamma}_{\alpha}),$$

$$\begin{aligned} \overset{(4)}{c}_{\alpha\beta} &= \frac{1}{4\,!}(2R_{0\alpha\beta0,00} + 8R^{\mu}_{0\alpha0}R_{0\mu\beta0} - 6R_{0\alpha\mu0,0}\Omega^{\mu}_{\beta} \\ &\quad - 6R_{0\beta\mu0,0}\Omega^{\mu}_{\alpha} + 8R_{0\mu\lambda0}\Omega^{\mu}_{\alpha}\Omega^{\lambda}_{\beta}), \end{aligned} \tag{8.2}$$

$$\begin{aligned} \overset{(5)}{c}_{\alpha\beta} &= \frac{1}{5\,!}\{2R_{0\alpha\beta0,000} + 14R^{\mu}_{0\alpha0}R_{0\mu\beta0,0} + 14R^{\mu}_{0\alpha0,0}R_{0\mu\beta0} \\ &\quad - 8\Omega^{\mu}_{\alpha}(R_{0\mu\beta0,00} + 2R^{\nu}_{0\mu0}R_{0\nu\beta0}) - 8\Omega^{\mu}_{\beta}(R_{0\mu\alpha0,00} \\ &\quad + 2R^{\nu}_{0\mu0}R_{0\nu\beta0}) + 20\Omega^{\lambda}_{\alpha}\Omega^{\mu}_{\beta}R_{0\lambda\mu0,0}\}. \end{aligned}$$

For $m=6$, we find

$$\begin{aligned} \overset{(6)}{c}_{\alpha\beta} &= \frac{1}{6\,!}(2R_{0\alpha\beta0,0000} + 22R^{\mu}_{0\alpha0}R_{0\mu\beta0,00} + 28R^{\mu}_{0\alpha0,0}R_{0\mu\beta0,0} \\ &\quad + 22R_{0\alpha0,00}R_{0\mu\beta0} + 32R^{\mu}_{0\nu0}R^{\nu}_{0\alpha0}R_{0\mu\beta0}) \\ &\quad - \frac{10}{6\,!}\Omega^{\mu}_{\alpha}(R_{0\mu\beta0,000} + 3R^{\nu}_{0\mu0}R_{0\nu\beta0,0} + 5R^{\nu}_{0\mu0,0}R_{0\nu\beta0}) \\ &\quad - \frac{10}{6\,!}\Omega^{\mu}_{\beta}(R_{0\mu\alpha0,000} + 3R^{\nu}_{0\mu0}R_{0\nu\alpha0,0} + 5R^{\nu}_{0\mu0,0}R_{0\nu\alpha0}) \\ &\quad + \frac{1}{6\,!}\Omega^{\lambda}_{\alpha}\Omega^{\mu}_{\beta}(36R_{0\lambda\mu0,00} + 32R^{\nu}_{0\lambda0}R_{0\nu\mu0}). \end{aligned} \tag{8.3}$$

Part III

9. We finally turn to the consideration of the tensors, and to obtaining some general properties. We have already observed in § 5 the geometrical interpretation Bianchi obtained for the functions $\Omega_{\alpha\beta}$. We shall here give another, one which has its analogue in 3-dimensional differential geometry.

Consider the hypersurface generated by the geodesics of V_{n+1} tangent to V_n at an arbitrary, but fixed point P of V_n; such a hypersurface we call the *tangent geodesic hypersurface* to V_n at P. We prove the following theorem:

---

* These values of $\overset{(m)}{c}_{\alpha\beta}$ were found by me first by making use of general coördinates. However, at the suggestion of G. Y. Rainich I made use of the particular coördinate system introduced in §4, and that not only enabled me to obtain my previous results more readily, but also made possible the determination of the general expression for $\overset{(m)}{c}_{\alpha\beta}$ (equations (7.10)).

If $\overline{V}_n$ is the tangent geodesic hypersurface to V_n at a point P and if P' is a point of $\overline{V}_n$ neighboring to P, then the distance from P' to V_n is given, except for terms of higher order, by $-\frac{1}{2}e\Omega_{\alpha\beta}du^\alpha du^\beta$.

The proof is very direct. Let us find a parametric representation of the tangent geodesic hypersurface. The geodesic through the point with coördinates $(u^{0}_{(0)}, u^{1}_{(0)}, \cdots, u^{n}_{(0)})$ and direction $(\zeta^0, \zeta^1, \cdots, \zeta^n)$ at this point has for its equation, when the coördinates of the space are the u's, the infinite series

$$(9.1)\quad u^i = u^i_{(0)} + \zeta^i s - \frac{1}{2}\left\{{i \atop rs}\right\}\zeta^r\zeta^s s^2 - \frac{1}{3!}\overline{\Gamma}^i_{rst}\zeta^r\zeta^s\zeta^t s^3 - \cdots \quad (i = 0,1, \cdots n),$$

where the functions $\overline{\Gamma}$ are formed with respect to the linear element (4.9) in precisely the way in which the Γ's of equations (2.10) are formed with respect to the linear element (1.1) and are evaluated for the point $u_{(0)}$. If we take for $u_{(0)}$ the point P with coördinates $(0, u^{1}_{(0)}, u^{2}_{(0)}, \cdots, u^{n}_{(0)})$ and for the direction one with components $(0, \zeta^1, \zeta^2, \cdots, \zeta^n)$, the geodesic is tangent to V_n at P. When we set $v^\alpha = \zeta^\alpha s$, equations (9.1) become

$$(9.2)\quad \begin{aligned} u^0 &= -\frac{1}{2}\left\{{0 \atop \beta\gamma}\right\}v^\beta v^\gamma - \frac{1}{3!}\overline{\Gamma}^0_{\beta\gamma\delta}v^\beta v^\gamma v^\delta - \cdots, \\ u^\alpha &= u^{\alpha}_{(0)} + v^\alpha - \frac{1}{2}\left\{{\alpha \atop \beta\gamma}\right\}v^\beta v^\gamma - \cdots. \end{aligned}$$

These equations, when we regard $v^1, v^2, \cdots, v^n$ as parameters, define the tangent geodesic hypersurface. Since the u^0 coördinate of a point $(u^0, u^1, \cdots, u^n)$ is the distance from the point to the hypersurface V_n, the required distance, D, from a point P' of the tangent geodesic hypersurface to V_n is given by

$$D = -\tfrac{1}{2}\left\{{0 \atop \beta\gamma}\right\}v^\beta v^\gamma.$$

But v^β is du^β except for terms of higher order; furthermore, by virtue of equations (4.8), $\left\{{0 \atop \beta\gamma}\right\} = e\Omega_{\beta\gamma}$. Hence

$$(9.3)\quad D = -\tfrac{1}{2}\, e\Omega_{\beta\gamma}du^\beta du^\gamma.$$

This proves the theorem. We may make the following further observations.

The contact between a hypersurface and its tangent geodesic hypersurface is in general simple. A necessary and sufficient condition that it be of order higher than the first at a point P and in the direction ζ is that

$$(9.4)\quad \Omega_{\alpha\beta}\zeta^\alpha\zeta^\beta = 0$$

at P. If (9.4) holds for every direction and at every point of the hypersurface, we have

$$\Omega_{\alpha\beta} = 0.$$

We observe that hypersurfaces for which $\Omega_{\alpha\beta}=0$ may be classified by means of the equations

$$\overset{(2)}{c}_{\alpha\beta} = \overset{(3)}{c}_{\alpha\beta} = \cdots = \overset{(m)}{c}_{\alpha\beta} = 0, \quad \overset{(m+1)}{c}_{\alpha\beta} \neq 0.$$

The larger the value of m for which this is true, the more nearly will the geometry of the space and the hypersurface be the same. In general a space does not contain any hypersurface for which $\overset{(m)}{c}_{\alpha\beta}=0$ for all m. We shall not at present carry the problems here suggested any further; instead we shall consider, in the remainder of this paper, a few special cases.

10. First suppose the form (1.1) or its equivalent (4.9) is reducible to a form with constant coefficients. Such a space we call a *flat* space. We observe that a flat space with a *positive definite* form is euclidean.

For a flat space, R_{ijkl} and its covariant derivatives vanish identically; from (6.13) it follows that $\overset{(m)}{c}_{\alpha\beta}=0$ for $m\geqq 3$. For $\overset{(m)}{c}_{\alpha\beta}$, $m<3$, equations (8.2) become

$$\overset{(0)}{c}_{\alpha\beta} = g_{\alpha\beta}, \quad \overset{(1)}{c}_{\alpha\beta} = -2\Omega_{\alpha\beta}, \quad \overset{(2)}{c}_{\alpha\beta} = \Omega_{\alpha\gamma}\Omega^{\gamma}_{\beta}. \tag{10.1}$$

Bianchi* has shown that

$$\Omega_{\alpha\delta} = -a_{ij}\frac{\partial\xi^i}{\partial u^\alpha}\frac{\partial f^j}{\partial u^\delta} - [ij,k]_a\frac{\partial f^i}{\partial u^\alpha}\frac{\partial f^k}{\partial u^\delta}\xi^j, \tag{10.2}$$

where the f's are those of equations (4.1), $[rs, t]_a$ are the Christoffel symbols of the space with reference to the form (1.1) and ξ^i is the vector normal to V_n referred to the general coördinate system of (1.1).

If the space is flat, we may take generalized cartesian coördinates, that is

$$\begin{aligned} a_{ij} &= 0 && (i \neq j), \\ a_{ii} &= 1 && (i = 0,1,\cdots,m), \\ a_{jj} &= -1 && (j = m+1,\cdots,n). \end{aligned} \tag{10.3}$$

Then we have

$$\Omega_{\alpha\delta} = \sum_{i=0}^{m}\frac{\partial\xi^i}{\partial u^\alpha}\frac{\partial f^i}{\partial u^\delta} - \sum_{j=m+1}^{N}\frac{\partial\xi^j}{\partial u^\alpha}\frac{\partial f^j}{\partial u^\delta}, \tag{10.4}$$

* Bianchi, loc. cit., p. 360.

and hence

$$c_{\alpha\beta}^{(2)} = \sum_{i=0}^{m} \frac{\partial \xi^i}{\partial u^\alpha}\frac{\partial \xi^i}{\partial u^\delta} - \sum_{j=m+1}^{N} \frac{\partial \xi^j}{\partial u^\alpha}\frac{\partial \xi^j}{\partial u^\delta}. \tag{10.5}$$

This is precisely the coefficient of the third fundamental form of the hypersurface for euclidean space* in which $m=n$.

If $\overline{V}_n$ is a hypersurface geodesically parallel to V_n and if we denote by $\bar{g}_{\alpha\beta}$, $\bar{\Omega}_{\alpha\beta}$, $\bar{c}_{\alpha\beta}^{(2)}$ the coefficients of its three fundamental forms, we have from equations (5.3) that

$$\begin{aligned} \bar{g}_{\alpha\beta} &= g_{\alpha\beta} - 2\Omega_{\alpha\beta}c + \Omega_{\alpha\gamma}\Omega^\gamma_\beta c^2, \\ \Omega_{\alpha\beta} &= \Omega_{\alpha\beta} - \Omega_{\alpha\gamma}\Omega^\gamma_\beta c, \\ \bar{c}_{\alpha\beta}^{(2)} &= \Omega_{\alpha\gamma}\Omega^\gamma_\beta, \end{aligned} \tag{10.6}$$

where c is the constant distance from $\overline{V}_n$ to V_n. We observe

In a flat space, hypersurfaces geodesically parallel have the same third fundamental form.

The last two theorems of § 12 hold for flat spaces but as the proof for the special case is similar to the proof for the general case we shall give only the latter in the later section.

11. Let us next consider the case when V_n is such that $R_{ijkl,m}=0$. This group of spaces contains as a subset the spaces of constant Riemann curvature; for if the curvature of V_{n+1} has the constant value of K_0, then†

$$R_{ijkl} = K_0(a_{ik}a_{jl} - a_{il}a_{jk}), \tag{11.1}$$

and from this it follows that $R_{ijkl,m}=0$. We can give a geometrical interpretation of the vanishing of $R_{ijkl,m}$ by means of the following theorems.

Let $\lambda^i (i=0, 1, 2, \cdots, n)$ be the components of an arbitrary contravariant vector, defining the congruence of curves whose differential equations are

$$\frac{dx^0}{\lambda^0} = \frac{dx^1}{\lambda^1} = \cdots = \frac{dx^n}{\lambda^n}.$$

Further, let $\lambda_{1|}{}^i$ and $\lambda_{2|}{}^i$ be the contravariant components of two vectors arbitrary except that each forms a family of parallels along and with respect to the curves of the given congruence, in the sense of Levi-Civita.

* Bianchi, loc. cit., p. 474; cf. Eisenhart, loc. cit., p. 219, for the general flat space.

† Bianchi, loc. cit., pp. 75, 343.

Then

A necessary and sufficient condition that the Riemannian curvature of (1.1) *be constant along the curves of the given congruence and for any pair of directions of the type of* $\lambda_{1|}$ *and* $\lambda_{2|}$ *is that* $R_{ijkl,m}\lambda^m$ *vanish.*

Before we prove this, let us observe that an immediate consequence of it is the following theorem:

If the curvature is constant along the curves of $n+1$* *independent congruences in arbitrary pairs of directions which remain parallel, then* $R_{ijkl,m}$ *vanishes.*

Let us return to the proof of the first theorem. The Riemann curvature K in the directions $\lambda_{1|}{}^i$ and $\lambda_{2|}{}^i$ when $\lambda_{1|}{}^i$ and $\lambda_{2|}{}^i$ are orthogonal unit vectors, is given by†

$$K = \pm R_{ijkl}\lambda_{1|}{}^i\lambda_{2|}{}^j\lambda_{1|}{}^k\lambda_{2|}{}^l. \tag{11.2}$$

If each of the directions $\lambda_{1|}$ and $\lambda_{2|}$ forms a system of parallels, in the sense of Levi-Civita, along the curves of a congruence defined by a vector λ^i we must have that‡

$$\lambda_{\epsilon|}{}^i{}_{,j}\lambda^j = 0 \qquad (\epsilon = 1,2\ ;\ i = 0,1, \cdots n) \tag{11.3}$$

where $\lambda_{\epsilon|}{}^i{}_{,j}$ is the covariant derivative of $\lambda_{\epsilon|}{}^i$ with respect to x^j and the fundamental form (1.1). Differentiate (11.2) covariantly with respect to x^m, multiply by λ^m and sum for m; because of (11.3) we obtain

$$\frac{\partial K}{\partial x^m}\lambda^m = R_{ijkl,m}\lambda_{1|}{}^i\lambda_{2|}{}^j\lambda_{1|}{}^k\lambda_{2|}{}^l\lambda^m. \tag{11.4}$$

If K is constant along the curves defined by λ^i, $(\partial K/\partial x^m)\lambda^m$ must vanish and since $\lambda_{1|}$ and $\lambda_{2|}$ are arbitrary, we have from (11.4) and (1.12) that

$$(R_{ijkl,m} + R_{kjil,m})\lambda^m = 0 \qquad (i,j,k,l = 0,1,2, \cdots, n).$$

Since this holds for all values of i, j, k and l, we have also

$$-(R_{iklj,m} + R_{lkij,m})\lambda^m = 0.$$

Furthermore, from (1.11) we have

$$(R_{ijkl,m} + R_{iklj,m} + R_{iljk,m})\lambda^m = 0.$$

* Note $n+1$ is the dimensionality of the space.

† Bianchi in deriving this formula does not assume that $\lambda_{1|}{}^i$ and $\lambda_{2|}{}^i$ are unit orthogonal vectors, so his value of K, loc. cit., p. 343, differs somewhat from (11.2). In what we do here, it is no restriction to make this hypothesis.

‡ Levi-Civita on Parallelism, Rendiconti del Circolo Matematico di Palermo, vol. 42 (1917), or Bianchi, Rendiconto della Reale Accademia delle Scienze Fisiche e Matematiche di Napoli, ser. 3, vol. 27 (1922).

Add these three equations, and we find

$$R_{ijkl,m}\lambda^m = 0. \tag{11.5}$$

Conversely, if (11.5) hold, from (11.4), $(\partial K/\partial x^m)\lambda^m$ vanishes, and K is constant along the curves defined by λ^i in directions parallel with respect to these curves. This proves the theorem.

For a space satisfying the condition of the second theorem* the sequence of tensors (8.2) becomes the following:

$$\begin{aligned}
c^{(2)}_{\alpha\beta} &= R_{0\alpha\beta 0} + \Omega_{\alpha\gamma}\Omega^{\lambda}_{\beta}, \\
c^{(3)}_{\alpha\beta} &= -\frac{4}{3\,!}\left(R_{0\alpha\delta 0}\Omega^{\delta}_{\beta} + R_{0\beta\delta 0}\Omega^{\delta}_{\alpha}\right), \\
c^{(2m)}_{\alpha\beta} &= \frac{2^{2m-1}}{(2m)\,!} R^{\lambda_1}_{0\lambda_2 0} R^{\lambda_2}_{0\lambda_3 0} \cdots R^{\lambda_{m-2}}_{0\lambda_{m-1}0}\left(R^{\lambda_{m-1}}_{0\alpha 0} R_{0\lambda_1\beta 0} + \Omega^{\lambda_{m-1}}_{\alpha}\Omega^{\mu}_{\beta} R_{0\lambda_1\mu 0}\right), \\
c^{(2m+1)}_{\alpha\beta} &= \frac{-2^{2m}}{(2m+1)\,!} R^{\lambda_1}_{0\lambda_2 0} R^{\lambda_2}_{0\lambda_3 0} \cdots R^{\lambda_{m-2}}_{0\lambda_{m-1}0} R^{\lambda_{m-1}}_{0\lambda_m 0}\left(R_{0\alpha\lambda_1 0}\Omega^{\lambda_m}_{\beta} + R_{0\beta\lambda_1 0}\Omega^{\lambda_m}_{\alpha}\right) \\
&\qquad (m = 2, 3, \cdots).
\end{aligned} \tag{11.6}$$

For the more restricted case of spaces of constant curvature equations (11.1) hold and in the coördinate system of (4.9) we have

$$R_{0\alpha\beta 0} = -eK_0 g_{\alpha\beta}, \qquad R^{\alpha}_{0\beta 0} = -eK_0\delta^{\alpha}_{\beta}.$$

Hence (11.6) become in this case

$$\begin{aligned}
c^{(2)}_{\alpha\beta} &= -eK_0 g_{\alpha\beta} + \Omega_{\alpha\gamma}\Omega^{\gamma}_{\beta}, \\
c^{(2m)}_{\alpha\beta} &= (-e)^{m-1}\frac{2^{2m-1}}{(2m)\,!} K_0^{m-1} c^{(2)}_{\alpha\beta}, \\
c^{(2m+1)}_{\alpha\beta} &= -(-e)^{m}\frac{2^{2m+1}}{2m\,!} K_0^{m}\Omega_{\alpha\beta} \qquad (m = 1, 2, \cdots).
\end{aligned} \tag{11.7}$$

The summation of (5.3) can be effected quite easily; we obtain

$$\begin{aligned}
c_{\alpha\beta} = g_{\alpha\beta} &- \frac{1}{\sqrt{eK_0}}\Omega_{\alpha\beta}\sin\left(2\sqrt{eK_0}\,u^0\right) \\
&+ \frac{1}{2eK_0}\left\{1 - \cos\left(2\sqrt{eK_0}\,u^0\right)\right\} c^{(2)}_{\alpha\beta}.\dagger
\end{aligned} \tag{11.8}$$

---

* That is, if $R_{ijkl,m}=0$.

† If eK_0 has a negative value, (11.8) must be replaced by a similar formula with hyperbolic functions instead of trigonometric.

Hence we may state the following theorem:

In a space of constant curvature, K_0, the fundamental tensor of a hypersurface geodesically parallel to and at a distance u^0 from an arbitrary hypersurface is given by (11.8).

By differentiation of (11.8) we obtain that for hypersurfaces geodesically parallel

$$
(11.9)\qquad
\begin{aligned}
\bar{\Omega}_{\alpha\beta} &= \Omega_{\alpha\beta}\cos(2\sqrt{eK_0}\,u^0) - \overset{(2)}{c}_{\alpha\beta}\,\frac{1}{\sqrt{eK_0}}\sin(2\sqrt{eK_0}\,u^0),\\
\bar{\overset{(2)}{c}}_{\alpha\beta} &= \Omega_{\alpha\beta}\;2\sqrt{eK_0}\sin(2\sqrt{eK_0}\,u^0) + \overset{(2)}{c}_{\alpha\beta}\cos(2\sqrt{eK_0}\,u^0).
\end{aligned}
$$

We shall return to the consideration of these equations at the end of the next section.

12. Let us finally consider a special type of hypersurfaces, which, in one sense, are analogous in the general space to spheres and planes in euclidean space. Bianchi has shown* that the curvature, $1/R$, of the geodesics of V is given by

$$
(12.1)\qquad \frac{1}{R} = \frac{\Omega_{\alpha\beta}du^\alpha du^\beta}{g_{\alpha\beta}du^\alpha du^\beta}.
$$

The maxima and minima of $1/R$ are given by the determinantal equation

$$
(12.2)\qquad \left|\,\Omega_{\alpha\beta} - \frac{1}{R}g_{\alpha\beta}\,\right| = 0,
$$

and the corresponding directions, called the *principal directions*, are solutions of

$$
(12.3)\qquad \left(\Omega_{\alpha\beta} - \frac{1}{R}g_{\alpha\beta}\right)\lambda^\beta = 0 \qquad (\alpha = 1,2,\cdots,n).
$$

There are n sets of principal directions corresponding to the n roots of (12.2); the n congruences of curves whose directions at every point coincide with the principal directions are the *lines of curvature*; $1/R$ is called the *normal curvature* of the hypersurface.

It follows immediately that a necessary and sufficient condition that the lines of curvature of a hypersurface be completely indeterminate is that $\Omega_{\alpha\beta} = (1/R)g_{\alpha\beta}$ for all α and β. These hypersurfaces are, in one sense, generalizations of spheres and planes; more particularly if $1/R = 0$ we have the type of hypersurface mentioned in § 9. We shall prove the following theorem:

* Bianchi, loc. cit., p. 366; Eisenhart, loc. cit., p. 151.

*In a space of constant Riemann curvature, the hypersurfaces whose lines of curvature are completely indeterminate have constant normal curvature and constant Riemann curvature; conversely if a hypersurface of space of constant Riemann curvature has constant Riemann curvature, then in the enveloping space it has constant normal curvature and completely indeterminate lines of curvature. The Riemann curvature of such a hypersurface is equal to the sum (or difference) of the square of its normal curvature and the Riemann curvature of the enveloping space.**

The proof is an immediate consequence of the equations of Gauss† connecting the first and second fundamental forms of a hypersurface. For spaces of constant curvature, K_0, these equations are

$$e(\Omega_{\alpha\gamma}\Omega_{\beta\delta} - \Omega_{\alpha\delta}\Omega_{\beta\gamma}) = \overline{R}_{\alpha\beta\gamma\delta} - K_0(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}), \tag{12.4}$$

where $\overline{R}_{\alpha\beta\gamma\delta}$ is the Riemann tensor formed with respect to the quadric form (4.3) and e is defined by (4.9). If the hypersurface has indeterminate lines of curvature, $\Omega_{\alpha\beta} = (1/R)g_{\alpha\beta}$ and consequently (12.4) become

$$\overline{R}_{\alpha\beta\gamma\delta} = \left(\frac{e}{R^2} + K_0\right)\left(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}\right). \tag{12.5}$$

From the theorem of Schur‡ it follows that $e/R^2 + K_0$ is constant, and from (11.1) it is the value of the Riemann curvature of V_n. Conversely, if (12.5) are satisfied, from (12.4) it follows that every two-row determinant from the square matrix $\|\Omega_{\alpha\beta}\|$ is equal to the corresponding two-row determinant from $\|g_{\alpha\beta}/R\|$. Furthermore, if $1/R \neq 0$ there is a three-row determinant from the latter matrix different from zero, since the determinant $|g_{\alpha\beta}| \neq 0$. It then follows from a theorem due to Killing§ that the elements of the two matrices differ at most in sign; that is

$$\Omega_{\alpha\beta} = \pm \frac{1}{R} g_{\alpha\beta} \qquad (\alpha, \beta = 1, 2, \cdots, n). \tag{12.6}$$

If $1/R = 0$, from (12.1) we have that $\Omega_{\alpha\beta} = 0$ and (12.6) hold in this case too. This proves the theorem.

We observe that e is necessarily positive if the form (1.1) or its equivalent (4.9) is positive definite; hence we have the following immediate corollary:

In spaces of more than three dimensions of constant Riemann curvature, K_0, whose fundamental form is positive definite, there are no real hypersurfaces of constant curvature $K < K_0$.

* The converse theorem holds only for $n > 2$; the direct theorem however is always true.

† Bianchi, loc. cit., p. 362; Eisenhart, loc. cit., p. 150.

‡ Mathematische Annalen, vol. 27, p. 563.

§ *Nicht-Euklidische Raumformen in Analytischer Behandlung*, Leipzig, Teubner, 1885, pp. 236–237

As a special case we have the well known theorem that in euclidean 4-space there are no 3-spaces of constant negative curvature.

If equations (12.6) hold, we find from (11.8), by virtue of the first of (11.7), that $c_{\alpha\beta}$ is proportional to $g_{\alpha\beta}$ and hence the map between V_n and $\overline{V}_n$ established by the geodesics of V_{n+1} normal to V_n is conformal. Conversely if a hypersurface is mapped conformally on all the hypersurfaces geodesically parallel to it, from (11.8) we obtain that $\Omega_{\alpha\beta}$ and $\overset{(2)}{c}_{\alpha\beta}$ are proportional to $g_{\alpha\beta}$. We observe further from (11.9) that, if $\Omega_{\alpha\beta}=(1/R)g_{\alpha\beta}$, we have that $\overline{\Omega}_{\alpha\beta}=(1/\overline{R})c_{\alpha\beta}$. Hence we have the following theorems:

If, in a space of constant curvature, a hypersurface is mapped conformally on the hypersurfaces geodesically parallel to it, the map being established by the normal geodesics, then these hypersurfaces have completely indeterminate lines of curvature.

The hypersurfaces geodesically parallel to a hypersurface with indeterminate lines of curvature in a space of constant Riemann curvature are themselves hypersurfaces with indeterminate lines of curvature.

Because of the first theorem in this section we may restate the preceding:

In a space of constant Riemann curvature, a hypersurface geodesically parallel to a hypersurface of constant Riemann curvature must also be of constant Riemann curvature.

In a flat space hypersurfaces with indeterminate lines of curvature must be hyperplanes or hyperspheres, and so the last two theorems are trivial.

From (12.3) we readily obtain that

$$\Omega_{\alpha\beta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta} = 0, \quad g_{\alpha\beta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta} = 0 \qquad (h,k = 1,2,\cdots,n;h \neq k), \tag{12.7}$$

and from (11.8) and (11.9) by virtue of (12.3) and (12.6) it follows that

$$c_{\alpha\beta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta} = 0,$$
$$\overline{\Omega}_{\alpha\beta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta} = 0 \qquad (h,k = 1,2,\cdots,n;h \neq k).$$

Equations (12.7) are also sufficient conditions that the congruences $\lambda_{h|}{}^{\alpha}$ $(h=1, 2, \cdots, n)$ be lines of curvature, for if in (12.7) we let h be fixed, and k take on the values $1, 2, \cdots, n, \neq h$ it follows that the vector $\Omega_{\alpha\beta}\lambda_{h|}{}^{\alpha}$ is normal to the $n-1$ vectors λ_k; hence $\Omega_{\alpha\beta}\lambda_{h|}{}^{\alpha}=\rho\lambda_{h|\beta}$ or $\lambda_{h|}$ is a line of curvature. Consequently we have the following theorem, which holds also in euclidean space:

In a space of constant curvature the lines of curvature of geodesically parallel hypersurfaces correspond.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

A COMPARISON OF THE SERIES OF FOURIER AND BIRKHOFF*

BY

M. H. STONE

In his thesis, Haar has made a comparative study of developments on $(0, \pi)$ in terms of the set $\sqrt{1/\pi}$, $\sqrt{2/\pi}\cos x$, $\sqrt{2/\pi}\cos 2x, \cdots$, and of a certain Sturm-Liouville set. He proves that the term-by-term difference of the formal series for any function summable with summable square converges uniformly to zero on the interval $(0, \pi)$†. Recently, Walsh, by entirely different and elegant methods, deduced a similar theorem for the set $\sqrt{2/\pi}\sin x$, $\sqrt{2/\pi}\sin 2x, \cdots$, and a second Sturm-Liouville set on the interval $(0, \pi)$‡. These two papers naturally suggested the possibility of analogous theorems for the series which Birkhoff defined in a memoir on the expansion problems of linear homogeneous differential systems of the nth order, and which are generalisations of the Sturm-Liouville series.§ Tamarkin, in a paper written to supplement the one just cited, compared the general Birkhoff series on the one hand and Fourier series on the other.|| He discussed only Riemann integrable functions and did not investigate the actual term-by-term difference of the series under consideration. In a later work on expansion problems which came to our attention after the completion of the present paper, he obtained, by methods different from our own, certain of our theorems, which will be noted subsequently.¶ Insofar as Tamarkin's treatise concerns us here, it consists in an extension of his previous results to the case of the summable function.

In the present paper we study the comparative properties not only of the Birkhoff and Fourier series for an arbitrary summable function, but also of the formal series obtained by deriving these term-by-term k times.

---

* Presented to the Society, December 30, 1924; accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Harvard University; received by the editors of these Transactions in October, 1925.

† Haar, Mathematische Annalen, vol. 69 (1910), pp. 331–371.

‡ Walsh, Annals of Mathematics, ser. 2, vol. 24, pp. 109–120.

§ Birkhoff, these Transactions, vol. 9(1908), pp. 373–395.

|| Tamarkin, Rendiconti del Circolo Matematico di Palermo, vol. 34 (1912), pp. 345–395.

¶ Tamarkin, *On Certain General Problems in the Theory of Ordinary Linear Differential Equations and the Expansion of an Arbitrary Function in Series*, Petrograd, 1917, in Russian: cited henceforth as D. E.

While our main purpose is the investigation of these questions, we first make generalisations of the series defined in Birkhoff's memoir. These generalisations are not trivial, and furthermore occupy a central position in the theory of linear differential equations. For this reason we take the liberty of describing them here.

I. Existence theorems

In the following pages we shall have constantly before us the differential equation

$$u^{(n)} + \bigstar + p_2 u^{(n-2)} + \cdots + p_{n-1}u^{(1)} + (p_n + \lambda)u = 0$$

where $p_2, \cdots, p_n$ are real or complex summable functions of the real variable x on the interval $(0, 1)$, and λ is a complex parameter. The usual existence theorems are obviously inapplicable to the present case. We employ a modification of the method of successive approximations to show the existence of "solutions" of such equations.* We first prove

Theorem I. *If*

(1) $p_1, \cdots, p_n, r$ *are real or complex summable functions of the real variable* x, $a \leqq x \leqq b$;

(2) λ *is a complex parameter restricted to the circle* $|\lambda| \leqq \Lambda$;

(3) x *is any point of* (a, b);

(4) $c_0, \cdots, c_{n-1}$ *are any real or complex constants*;

then there exists a constant $\delta_\Lambda > 0$, *independent of* $\lambda, x_0, c_0, \cdots, c_{n-1}$; *and there exists a function* $u(x, \lambda)$, *defined for all* λ, $|\lambda| \leqq \Lambda$, *and for all* x *on the interval common to* (a, b) *and* $(x_0 - \delta_\Lambda, x_0 + \delta_\Lambda)$, *such that*

(1) $u^{(k)}(x, \lambda)$ *is continuous in* x *and analytic in* λ, $k = 0, \cdots, n-1$, *the differentiation being with respect to* x;

(2) $u^{(k)}(x_0, \lambda) = c_k$, $k = 0, \cdots, n-1$;

(3) $u^{(n)}(x, \lambda)$ *is summable in* x *for each value of* λ;

(4) $u^{(n-1)}(x, \lambda) = c_{n-1} + \int_{x_0}^{x} u^{(n)}(x, \lambda)dx$;

(5) $u^{(n)} + p_1 u^{(n-1)} + \cdots + p_{n-1}u^{(1)} + (p_n + \lambda) = r$, *except possibly on a set of zero measure*, E_λ, *for each* λ;

(6) $u(x, \lambda)$ *is unique.*

In the case $p_1 = p_2 = \cdots = p_n = 0$, $\Lambda = 0$, the equation reduces to $u^{(n)} = r$; and we can construct a function with the properties (1)–(5) of the theorem for the whole interval (a, b). By direct verification we see that it is

$$u = c_0 + \frac{c_1(x - x_0)}{2} + \cdots + \frac{c_{n-1}(x - x_0)^{n-1}}{(n-1)!} + \int_{x_0}^{x} \frac{(x - \xi)^{n-1}}{(n-1)!} r(\xi)d\xi.$$

---

* Tamarkin has obtained existence theorems here differently, D. E., Chapter I.

Hence in the general case we attempt to solve the equation

$$u = c_0 + \frac{c_1(x - x_0)}{2} + \cdots + \frac{c_{n-1}(x - x_0)^{n-1}}{(n-1)!} + \int_{x_0}^{x} \frac{(x-\xi)^{n-1}}{(n-1)!} [S(u) + r] d\xi,$$

$$S(u) = -p_1 u^{(n-1)} - \cdots - p_{n-1} u^{(1)} - (p_n + \lambda) u.$$

We first form the sequence of successive approximants $[u_i]$ where

$$u_0 = c_0 + \frac{c_1(x - x_0)}{2} + \cdots + \frac{c_{n-1}(x - x_0)^{n-1}}{(n-1)!},$$

$$u_{i+1} = u_0 + \int_{x_0}^{x} \frac{(x-\xi)^{n-1}}{(n-1)!} [S(u_i) + r(\xi)] d\xi.$$

To consider the convergence of this sequence we construct a second sequence where

$$v_0 = u_0,$$

$$v_1 = u_1 - u_0 = \int_{x_0}^{x} \frac{(x-\xi)^{n-1}}{(n-1)!} [S(u_0) + r] d\xi,$$

$$v_{i+1} = u_{i+1} - u_i = \int_{x_0}^{x} \frac{(x-\xi)^{n-1}}{(n-1)!} S(v_i)\, d\xi \qquad (i = 1, 2, \cdots).$$

Clearly, for $k = 0, 1, 2, \cdots, n-1$ and $i = 1, 2, \cdots$, and for all x, x_0, ξ on (a, b) we have

$$|v_1^{(k)}| \leqq C, \qquad \left| \frac{(x-\xi)^{(n-k-1)}}{(n-k-1)!} \right| \leqq L,$$

and

$$v_{i+1}^{(k)} = \int_{x_0}^{x} \frac{(x-\xi)^{n-k-1}}{(n-k-1)!} S(v_i) d\xi.$$

Next, by a fundamental property of the Lebesgue integral, we can write

$$\left| \int_{x_0}^{x} (|p_1| + \cdots + |p_n| + |\lambda|) d\xi \right| \leqq \frac{1}{2L}$$

for all λ, $|\lambda| \leqq \Lambda$, and for all x, $|x - x_0| \leqq \delta_\Lambda$, *where* $\delta_\Lambda > 0$ *does not depend on* x_0. We now show that if the inequality $|v_i^{(k)}| \leqq C/2^{i-1}$ is true, $|x - x_0| \leqq \delta_\Lambda$, it is also true when i is replaced by $i+1$; since it holds for $i = 1$, it will be true for all i. We have

$$|v_{i+1}^{(k)}| \leqq L \left| \int_{x_0}^{x} (|p_1| + \cdots + |p_n| + |\lambda|) d\xi \right| \frac{C}{2^{i-1}} \leqq \frac{C}{2^i},$$

$$|x - x_0| \leqq \delta_\Lambda.$$

Hence the infinite series $v_0^{(k)}+v_1^{(k)}+v_2^{(k)}+\cdots$, $k=0,\cdots,n-1$, converge uniformly, $|x-x_0|\leqq\delta_\Lambda$, $|\lambda|\leqq\Lambda$. The individual terms are continuous in x and polynomial in λ. We can thus be sure that the limit functions are continuous in x and analytic in λ, and can write

$$\lim_{i\to\infty}\sum_{\alpha=1}^{\alpha=i} v_\alpha = \lim_{i\to\infty} u_i^{(k)} = u^{(k)}(x,\lambda).$$

The functions $u^{(k)}(x,\lambda)$ therefore satisfy conclusion (1) of our theorem. On allowing i to become infinite in the relation connecting u_{i+1} and u_i, we find

$$u(x,\lambda) = u_0(x) + \int_{x_0}^{x} \frac{(x-\xi)^{n-1}}{(n-1)!}[S(u)+r]d\xi, \quad |x-x_0|\leqq\delta_\Lambda.$$

By direct computation we see that (2) is fulfilled. On deriving the identity n times with respect to x there results

$$u^{(n)}(x,\lambda) = S(u) + r(x)$$

except possibly on a set E_λ of measure zero. Thus (3) and (5) are satisfied. Then we see that

$$\begin{aligned} u^{(n-1)}(x,\lambda) &= c_{n-1} + \int_{x_0}^{x}[S(u)+r]d\xi \\ &= c_{n-1} + \int_{x_0}^{x} u^{(n)}(x,\lambda)dx \end{aligned}$$

which is conclusion (4).

We now assume that on some interval containing x there exists a second function $\bar{u}(x,\lambda)$, $|\lambda|\leqq\Lambda$, with the properties

(1′) $\bar{u}^{(k)}(x,\lambda)$, $k=0,\cdots,n-1$, is continuous in x for each value of λ;
(2′) $\bar{u}^{(k)}(x_0,\lambda)=c_k$, $k=0,\cdots,n-1$;
(3′) $\bar{u}^{(n)}(x,\lambda)$ is summable in x for each value of λ;
(4′) $\bar{u}^{(n-1)}(x,\lambda)=c_{n-1}+\int_{x_0}^{x}\bar{u}^{(n)}(x,\lambda)dx$;
(5′) $\bar{u}^{(n)}+p_1\bar{u}^{(n-1)}+\cdots+p_{n-1}\bar{u}^{(1)}+(p_n+\lambda)\bar{u}=r$, except on a set E_λ' of zero measure.

The difference $U=u-\bar{u}$ is then a function satisfying the conditions (1′)-(5′) with $c_k=0$, $k=0,\cdots,n-1$, and $r(x)\equiv 0$, the set E_λ' being replaced by the set $E_\lambda''=E_\lambda+E_\lambda'$. We suppose that at some point $(\bar{x}_0,\lambda_0)$ in the domain of definition of $U(x,\lambda)$ we have $U(\bar{x}_0,\lambda_0)\neq 0$.

By the part of the theorem already proved we now construct functions $U_2,\cdots,U_n$, satisfying the conditions (1)-(5) with $U_i^{(k)}(\bar{x}_0,\lambda)=\delta_{i,k+1}$, the well known Kronecker symbol, and $r(x)\equiv 0$. We then form

$$W(x,\lambda) = \begin{vmatrix} U & U' & \cdots & U^{(n-1)} \\ U_2 & U_2' & \cdots & U_2^{(n-1)} \\ \cdot & \cdot & \cdots & \cdot \\ U_n & U_n' & \cdots & U_n^{(n-1)} \end{vmatrix}$$

with $W(x_0, \lambda)=0$, $W(\bar{x}_0, \lambda)=U(\bar{x}_0, \lambda)\neq 0$. It is easily seen by the use of (5) and (5′) that we have

$$W'(x,\lambda) = -p_1(x)W(x,\lambda)$$

almost everywhere, while by (3) and (3′) it is clear that

$$W(x,\lambda) = \int_{x_0}^{x} W'(x,\lambda)dx.$$

On the other hand the function

$$W_0(x) = e^{-\int_{x_0}^{x} p_1 dx}$$

satisfies

$$W_0'(x) = -p_1(x)W_0(x)$$

almost everywhere, and

$$W_0(x) = 1 + \int_{x_0}^{x} W_0'(x)dx.$$

We then find that $(W/W_0)'=0$ almost everywhere; and we can integrate this relation obtaining $W=c(\lambda)W_0$. Since $W(x_0, \lambda)=0$, it follows that $c(\lambda)=0$, $W(x, \lambda)\equiv 0$. This leads to the contradictory statement $W(\bar{x}_0, \lambda) = U(\bar{x}_0, \lambda)=0$. We thus have $U(x, \lambda)\equiv 0$ over the domain for which it is defined; in other words, $u(x, \lambda)$ is unique, as asserted in the sixth point of the theorem.

It is now easy to demonstrate the general existence theorem, which is

THEOREM II. *Under the hypotheses* (1)–(4) *of Theorem* I *there exists a unique function* $u(x, \lambda)$, $a\leqq x\leqq b$, *such that*

(1) $u^{(k)}(x, \lambda)$ *is continuous in x and entire in* λ, $k=0, \cdots, n-1$;

(2) $u^{(k)}(x_0, \lambda)=c_k$, $k=0, \cdots, n-1$;

(3) $u^{(n)}(x, \lambda)$ *is summable in x for each value of* λ;

(4) $u^{(n-1)}(x, \lambda)=c_{n-1}+\int_{x_0}^{x}u^{(n)}(x, \lambda)dx$;

(5) $u^{(n)}+p_1u^{(n-1)}+\cdots+p_{n-1}u^{(1)}+(p_n+\lambda)u=r$, *except possibly on a set* $\mathcal{E}$ *of measure zero, independent of* λ.

At x_0 we form the interval I_0: $|x-x_0|\leqq\delta_\Lambda$; at the right-hand end of I_0, $x_0+\delta_\Lambda$, we form the interval I_1: $|x-x_0-\delta_\Lambda|\leqq\delta_\Lambda$; and we continue, forming

$I_2, \cdots, I_N$, where N is large enough that I_N contains $x=b$. Similarly, we form a chain of intervals $I_{-1}, \cdots, I_{-M}$ running to the left, with M large enough that I_{-M} contains $x=a$. Then in I_0 we use Theorem I to set up a function $u(x, \lambda)$ satisfying conditions (1)–(6) of that theorem. In I_1 we apply Theorem I to define a function $u_1(x, \lambda)$ satisfying (1)–(6) in I_1, with $u_1^{(k)}(x_0+\delta_\Lambda, \lambda)=u^{(k)}(x_0+\delta_\Lambda, \lambda)$; by (6), u and u_1 coincide, $x_0 \leqq x \leqq x_0+\delta_\Lambda$, $|\lambda| \leqq \Lambda$. We then perform a similar set of operations in $I_2, \cdots, I_N, I_{-1}, \cdots, I_{-M}$. The result is the formation of the function $u(x, \lambda)$ whose existence is asserted in the present theorem.

For this function properties (1)–(4) are immediate. In order to prove (5) we start from the fact that for each λ the differential equation is satisfied except on a set E_λ of zero measure. We denote by λ' the points of the λ-plane with rational coördinates. We then define the set of zero measure

$$\mathcal{E}_1 = \sum_{(\lambda')} E_{\lambda'}.$$

Writing

$$P(x) = \int_{x_0}^{x} (|p_1| + \cdots + |p_n| + 1)dx,$$

we denote by $\mathcal{E}_2$ the set of zero measure on which P' does not exist, is infinite or does not coincide with the integrand in the definition of P. We show that we can set $\mathcal{E}=\mathcal{E}_1+\mathcal{E}_2$. We let x' be any point of (a, b) not in $\mathcal{E}$, λ any point of the complex plane, and λ' a point sufficiently close to λ that for all x on (a, b) and for $k=0, \cdots, n-1$, the inequalities

$$|u^{(k)}(x,\lambda) - u^{(k)}(x,\lambda')| < \epsilon, \quad |\lambda u(x,\lambda) - \lambda' u(x,\lambda')| < \epsilon$$

are true. We find

$$\begin{aligned} &\limsup_{h\to 0} \left| \frac{1}{h} u^{(n-1)}(x,\lambda)\Big]_{x=x'}^{x=x'+h} - \frac{1}{h} u^{(n-1)}(x,\lambda')\Big]_{x=x'}^{x=x'+h} \right| \\ &= \limsup_{h\to 0} \left| \frac{1}{h} \int_{x'}^{x'+h} \left\{ p_1 u^{(n-1)}\Big]_{\lambda}^{\lambda'} + \cdots + p^n u\Big]_{\lambda}^{\lambda'} + \lambda' u(x,\lambda') - \lambda u(x,\lambda) \right\} dx \right| \\ &\leqq \epsilon \lim_{h\to 0} \frac{P(x'+h) - P(x')}{h} = \epsilon P'(x'). \end{aligned}$$

Since

$$\lim_{h\to 0} \frac{1}{h} u^{(n-1)}(x,\lambda')\Big]_{x=x'}^{x=x'+h} = -p_1 u^{(n-1)} - \cdots - p_n u - \lambda' u + r,$$

and since ϵ can be taken arbitrarily small, we must have

$$\lim_{h\to\infty} \frac{1}{h} u^{(n-1)}(x,\lambda)\Big]_{x=x'}^{x=x'+h} = -p_1 u^{(n-1)} - \cdots - p_n u - \lambda u + r,$$

as we were to show.

All the usual discussions of linear dependence and independence of solutions, solutions satisfying given boundary conditions, Green's function, and so forth, can now be repeated for the general equation described in the theorem.* We may note that a consideration of the linear dependence of solutions would lead to a sharpening of Theorem II in that the set $\mathcal{E}$ could be replaced by a set E, independent of the particular solution.

We turn next to another type of theorem. While that which we shall now discuss differs but slightly from a less general theorem of Birkhoff in its proof, we take the opportunity of giving a complete demonstration since the details of the proof can be simplified for our purposes.†

We denote by S a sector of the type $l\pi/n \leqq \theta \leqq (l+1)\pi/n$, $\arg \rho = \theta$, where $l=0, 1, 2, \cdots, 2n-1$. From the sector S we define a region T by an arbitrary translation taking the vertex into the point $-c$, where c is a complex constant. The n distinct roots of $\omega^n+1=0$ can be written in the order $\omega_1, \cdots, \omega_n$ such that for all ρ on T

$$\Re(\overline{\rho + c\omega_1}) \leqq \Re(\overline{\rho + c\omega_2}) \leqq \cdots \leqq \Re(\overline{\rho + c\omega_n})$$

where the notation $\Re(\ )$ means, as usual, "the real part of".‡ We commence with two important lemmas. The first is

Lemma I. *On the region T of the ρ-plane*

$$\left| \frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} \right| \leqq C \,|\rho|^k i \,|e^{\rho\omega_i(x-\xi)}|, \qquad 0 \leqq \xi \leqq x \leqq 1,$$

$$\left| \frac{d^k}{dx^k} \sum_{\alpha=i+1}^{\alpha=n} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} \right| \leqq C \,|\rho|^k (n-i) \,|e^{\rho\omega_i(x-\xi)}|, \quad 0 \leqq x \leqq \xi \leqq 1,$$

for $k=0, \cdots, n-1$, $i=1, \cdots, n$. The constant C may be taken independent of i and k.

We let C be chosen so that

$$|e^{c(\omega_i-\omega_\alpha)(x-\xi)}| \leqq C, \ \alpha = 1, \cdots, i, i = 1, \cdots, n, 0 \leqq \xi \leqq x \leqq 1,$$
$$|e^{c(\omega_i-\omega_\alpha)(x-\xi)}| \leqq C, \ \alpha = i+1, \cdots, n, i = 1, \cdots, n, 0 \leqq x \leqq \xi \leqq 1.$$

Then we have for $\alpha \leqq i$ and for ρ on T

$$\Re(\rho\omega_\alpha) \leqq \Re(\rho\omega_i + \overline{c\omega_i - \omega_\alpha}).$$

* For the usual existence theorems and subsequent developments, see Bôcher, *Leçons sur les Méthodes de Sturm*, Paris, 1917.

† Birkhoff, these Transactions, vol. 9 (1908), pp. 219–231.

‡ Birkhoff, these Transactions, vol. 9 (1908), p. 381.

Thus, if $x \geqq \xi$,

$$\left| \frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} \right| \leqq |\rho|^k \sum_{\alpha=1}^{\alpha=i} \left| e^{\rho\omega_\alpha(x-\xi)} \right| \leqq |\rho|^k \sum_{\alpha=1}^{\alpha=i} \left| e^{(\rho\omega_i + c\overline{\omega_i - \omega_\alpha})(x-\xi)} \right|$$

$$\leqq C |\rho|^k i \left| e^{\rho\omega_i(x-\xi)} \right|, \quad 0 \leqq \xi \leqq x \leqq 1,$$

as we were to show.

Again, starting from the inequality $\mathcal{R}(\rho\omega_\alpha) \geqq \mathcal{R}(\rho\omega_i + c\overline{\omega_i - \omega_\alpha})$, $\alpha \geqq i$, which is true on T, we derive the second inequality by analogous means.

In Lemma II we shall employ heavy-faced type to denote square matrices of n^2 elements.* The statement of the lemma is as follows:

Lemma II. *If* $\mathbf{\Omega}(x)$ *has all its elements real or complex continuous functions of* x, $0 \leqq x \leqq 1$, *and if* $\mathbf{K}(x, \xi, \rho)$ *has all its elements of the form* $p(\xi)k(x, \xi, \rho)$ *where*

(1) $p(\xi)$ *is a real or complex summable function,* $0 \leqq \xi \leqq 1$;

(2) $k(x, \xi, \rho)$ *is continuous in* x *and* ξ *for* $0 \leqq \xi < x \leqq 1$ *and for* $0 \leqq x < \xi \leqq 1$, *and is analytic in* ρ, $|\rho| \geqq R$ *on* T;

(3) $|k(x, \xi, \rho)| \leqq k$, $0 \leqq x \leqq 1$, $0 \leqq \xi \leqq 1$, $|\rho| \geqq R$ *on* T;

then there exists a matrix $\mathbf{Z}(x, \rho)$, $0 \leqq x \leqq 1$, $|\rho|$ *sufficiently large on* T, *whose elements are continuous in* x *and analytic in* ρ, *and which satisfies*

$$\mathbf{Z}(x,\rho) = \mathbf{\Omega}(x) + \frac{1}{\rho}\int_0^1 \mathbf{Z}(\xi,\rho)\mathbf{K}(x,\xi,\rho)d\xi.$$

Furthermore,

$$\mathbf{Z}(x,\rho) = \mathbf{\Omega}(x) + \frac{1}{\rho}\mathbf{E}(x,\rho) \text{ where } |\mathbf{E}(x,\rho)| \leqq \mathbf{M},\ 0 \leqq x \leqq 1,$$

ρ *on* T.

If $\mathbf{U}$ denotes the matrix all of whose elements are unity, we can write

$$|\mathbf{\Omega}| \leqq \omega \mathbf{U}, \int_0^1 |\mathbf{K}(x,\xi,\rho)| d\xi \leqq p\mathbf{U}.$$

We consider the infinite series

$$\mathbf{Z}(x,\rho) = \mathbf{\Omega}(x) + \sum_{k=1}^{\infty} \frac{1}{\rho^k}\int_0^1 \cdots \int_0^1 \Omega(\xi_{k-1})\mathbf{K}(\xi_{k-2},\xi_{k-1},\rho) \cdots \mathbf{K}(x,\xi,\rho)d\xi_{k-1} \cdots d\xi.$$

---

* For further details of notation see Birkhoffa nd Langer, Proceedings of the American Academy of Arts and Sciences, vol. 58 (1923), pp. 52–54.

We see that the general term is in absolute value less than or equal to

$$\frac{1}{|\rho|^k}\omega p^k \boldsymbol{U}^{k+1} = \frac{n^{k+1}\omega p^k}{|\rho|^k}\boldsymbol{U}.$$

Thus if we take $np/|\rho| \leqq \theta < 1$, the series converges absolutely and uniformly in x and ρ. The individual terms are continuous in x and analytic in ρ; hence this is true of the elements of $\boldsymbol{Z}$. By direct computation with this series we find that

$$\boldsymbol{Z}(x,\rho) = \boldsymbol{\Omega}(x) + \frac{1}{\rho}\int_0^1 \boldsymbol{Z}(\xi,\rho)\boldsymbol{K}(x,\xi,\rho)d\xi.$$

We now let $m(\rho)$ be the maximum value taken on by the elements of $\boldsymbol{Z}(x, \rho)$ for $0 \leqq x \leqq 1$. Then we have $|\boldsymbol{Z}(x, \rho)| \leqq m(\rho)\boldsymbol{U}$. From the integral equation

$$|\boldsymbol{Z}(x,\rho)| \leqq \omega\boldsymbol{U} + \frac{1}{|\rho|}pm(\rho)\boldsymbol{U}^2 = \left(\omega + \frac{pnm(\rho)}{|\rho|}\right)\boldsymbol{U}.$$

All the elements of the matrix on the right are equal; there is always some element of the matrix on the left which for an appropriate value of x takes on the value $m(\rho)$. Hence

$$m(\rho) \leqq \omega + \frac{npm(\rho)}{|\rho|},\quad m(\rho) \leqq \frac{\omega}{1 - \dfrac{pn}{|\rho|}} \leqq m, \quad |\rho| \geqq R'.$$

Now we find the inequality

$$|\boldsymbol{Z}(x,\rho) - \boldsymbol{\Omega}(x)| \leqq \frac{1}{|\rho|}m(\rho)p\boldsymbol{U}^2 \leqq \frac{nmp}{|\rho|}\boldsymbol{U}$$

from which the rest of the lemma follows.

We come now to

THEOREM III. *If $p_2, \cdots, p_n$ are real or complex summable functions of the real variable x, $0 \leqq x \leqq 1$, and if ρ is a complex parameter, then the differential equation $u^{(n)} + \bigstar + p_2u^{(n-2)} + \cdots + p_{n-1}u^{(1)} + (p_n + \rho^n)u = 0$ has on any region T of the complex ρ-plane, n linearly independent solutions analytic in ρ, $u_1, \cdots, u_n$, expressible with their first $n-1$ derivatives in the form*

$$u_i^{(k)}(x,\rho) = \rho^k e^{\rho\omega_i x}\left(\omega_i^k + \frac{E_{ik}(x,\rho)}{\rho}\right)(i = 1, \cdots, n;\ k = 0, \cdots, n-1),$$

where $|E_{ik}(x, \rho)| \leqq M$ for all ρ on T, $|\rho| \geqq R$, and for all x, $0 \leqq x \leqq 1$.

We write the differential equation as

$$u^{(n)} + \rho^n u = - p_2 u^{(n-2)} - \cdots - p_n u = - S'(u),$$

and proceed to solve this equation as though the right-hand side were known. Thus

$$u = c_1 e^{\rho\omega_1 x} + \cdots + c_n e^{\rho\omega_n x} + \int_0^x \frac{\omega_1 e^{\rho\omega_1(x-\xi)} + \cdots + \omega_n e^{\rho\omega_n(x-\xi)}}{n\rho^{n-1}} S'(u)d\xi,$$

where $c_1, \cdots, c_n$ are arbitrary constants. We now write

$$\bar{c}_l = c_l \qquad (l = 1, \cdots, i),$$

$$\bar{c}_l = c_l + \sum_{\alpha=i+1}^{\alpha=n} \int_0^1 \frac{\omega_\alpha e^{-\rho\omega_\alpha \xi}}{n\rho^{n-1}} S'(u)d\xi \qquad (l = i + 1, \cdots, n).$$

We find at once

$$u = \bar{c}_1 e^{\rho\omega_1 x} + \cdots + \bar{c}_n e^{\rho\omega_n x} + \frac{1}{n\rho^{n-1}}\int_0^x \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} S'(u)d\xi$$
$$+ \frac{1}{n\rho^{n-1}}\int_1^x \sum_{\alpha=i+1}^{\alpha=n} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} S'(u)d\xi.$$

Next we let u_i correspond to $\bar{c}_1 = \cdots = \bar{c}_{i-1} = \bar{c}_{i+1} = \cdots = \bar{c}_n = 0$, $\bar{c}_i = 1$. We then put $u_i^{(k)} = \rho^k e^{\rho\omega_i x} z_{ik}$ for $i = 1, \cdots, n$ and $k = 0, \cdots, n-1$. Then it is apparent that

$$\rho^k e^{\rho\omega_i x} z_{ik} = \omega_i^k \rho^k e^{\rho\omega_i x} + \frac{1}{n\rho^{n-1}}\int_0^x \left\{\frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\} S'(u)d\xi$$
$$+ \frac{1}{n\rho^{n-1}}\int_1^x \left\{\frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=n} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\} S'(u)d\xi.$$

On expressing $S'(u)$ in terms of the z_{ik} this equation becomes

$$z_{ik} = \omega_i^k + \frac{1}{n\rho}\int_0^x e^{-\rho\omega_i(x-\xi)}\rho^{-k}\left\{\frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\}\left\{p_2 z_{i,n-2} + \cdots + \frac{p_n}{\rho^{n-2}} z_{i,0}\right\} d\xi$$
$$+ \frac{1}{n\rho}\int_1^x e^{-\rho\omega_i(x-\xi)}\rho^{-k}\left\{\frac{d^k}{dx^k} \sum_{\alpha=i+1}^{\alpha=1} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\}\left\{p_2 z_{i,n-2} + \cdots + \frac{p_n}{\rho^{n-2}} z_{i,0}\right\} d\xi.$$

By Lemma I we see that this set of equations is precisely of the sort discussed in Lemma II. Thus we obtain solutions $u_1, \cdots, u_n$ such that $u_i^{(k)}(x, \rho)$ is continuous in x, $0 \leqq x \leqq 1$, and analytic in ρ, $|\rho| \geqq R$, on T, for $i=1, \cdots, n$, $k=0, \cdots, n-1$, while

$$u_i^{(k)}(x,\rho) = \rho^k e^{\rho\omega_i x}\left\{\omega_i^k + \frac{E_{ik}(x,\rho)}{\rho}\right\}, \qquad |E_{ik}| \leqq M.$$

The functions $u_1, \cdots, u_n$ are clearly linearly independent for large $|\rho|$, by their asymptotic form.

It remains for us to show that the $u_1, \cdots, u_n$ satisfy the differential equation. We know that there exist solutions $\bar{u}_1, \cdots, \bar{u}_n$ of the differential equation, linearly independent for all values of ρ; we need only apply Theorem II taking $\bar{u}_i^{(k)}(0) = \delta_{i,k+1}$, $i=1, \cdots, n$, $k=0, \cdots, n-1$. By assigning appropriate values, analytic in ρ, to the coefficients $\bar{c}_1, \cdots, \bar{c}_n$ of the integral equation above, we obtain particular equations satisfied by $\bar{u}_1, \cdots, \bar{u}_n$. We show that any solution u for a similar equation in which $\bar{c}_1, \cdots, \bar{c}_n$ are constants can be expressed linearly in terms of $\bar{u}_1, \cdots, \bar{u}_n$ and is consequently a solution of the differential equation. We write

$$U = u - c_1\bar{u}_1 - \cdots - c_n\bar{u}_n,$$

$$U = A_1 e^{\rho\omega_1 x} + \cdots + A_n e^{\rho\omega_n x} + \frac{1}{n\rho^{n-1}} \int_0^x \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} S'(U) d\xi$$

$$+ \frac{1}{n\rho^{n-1}} \int_1^x \sum_{\alpha=i+1}^{\alpha=n} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)} S'(U) d\xi.$$

We see now that either (1) when the constants $\bar{c}_1, \cdots, \bar{c}_n$ are not all zero, we can determine $A_1, \cdots, A_n$ identically zero by an appropriate choice of $C_1, \cdots, C_n$; or (2) we can take $u=0$ and determine $C_1, \cdots, C_n$ not all zero so that $A_1, \cdots, A_n$ are identically zero. Either possibility leads us to study the equation last written down with the coefficients $A_1, \cdots, A_n$ taken equal to zero. On writing $U^{(k)} = \rho^k e^{\rho\omega_i x} Z_k$ we find

$$Z_k = \frac{1}{n\rho^{n-1}} \int_0^x \rho^{-k} e^{-\rho\omega_i(x-\xi)} \left\{\frac{d^k}{dx^k} \sum_{\alpha=1}^{\alpha=i} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\} \{p_2\rho^{n-2} Z_{n-2} +$$

$$\cdots + p_n Z_0\} d\xi$$

$$+ \frac{1}{n\rho^{n-1}} \int_1^x \rho^{-k} e^{-\rho\omega_i(x-\xi)} \left\{\frac{d^k}{dx^k} \sum_{\alpha=i+1}^{\alpha=n} \omega_\alpha e^{\rho\omega_\alpha(x-\xi)}\right\} \{p_2\rho^{n-2} Z_{n-2} +$$

$$\cdots + p_n Z_0\} d\xi.$$

If $m(\rho)$ is the maximum value taken on by $|Z_k(x, \rho)|$, $0 \leqq x \leqq 1$, $k=0, \cdots, n-1$, Lemma I enables us to conclude that

$$|Z_k| \leqq \frac{1}{n|\rho|} Ci \int_0^x \left\{ |p_2| + \cdots + \frac{|p_n|}{|\rho|^{n-2}} \right\} d\xi \cdot m(\rho)$$
$$+ \frac{1}{n|\rho|} C(n-i) \int_x^1 \left\{ |p_2| + \cdots + \frac{|p_n|}{|\rho|^{n-2}} \right\} d\xi \cdot m(\rho).$$

For some values of x and k we can always write $|Z_k| = m(\rho)$, so that

$$m(\rho) \leqq m(\rho) \frac{C}{|\rho|} \int_0^1 \left\{ |p_2| + \cdots + \frac{|p_n|}{|\rho|^{n-2}} \right\} d\xi \leqq m(\rho) \frac{K}{|\rho|}.$$

This is impossible for large values of $|\rho|$ unless $m(\rho)=0$. Thus $U\equiv 0$. Under the second of the two alternatives above we should have $C_1\bar{u}_1 + \cdots + C_n\bar{u}_n$ identically zero without having $C_1, \cdots, C_n$ all identically zero. By hypothesis this is impossible. The first alternative alone is possible. Consequently we have $u \equiv C_1\bar{u}_1 + \cdots + C_n\bar{u}_n$ as we wished to show. We may remark that, since the coefficients $\bar{c}_1, \cdots, \bar{c}_n$ associated with u were assumed to be constants, we can assert that $C_1, \cdots, C_n$ are analytic for all ρ of sufficiently great absolute value so that u is also analytic in ρ. This completes the theorem.

We shall need in some of our work a slightly more detailed form of the theorem actually stated by Birkhoff.

THEOREM III′. *If $p_2, \cdots, p_n$ are real or complex functions of the real variable x, $0 \leqq x \leqq 1$, continuous together with their derivatives of all orders, and if ρ is a complex parameter, then the differential equation*

$$u^{(n)} + \bigstar + p_2 u^{(n-2)} + \cdots + p_{n-1}u^{(1)} + (p_n + \rho^n)u = 0$$

has on any region T of the complex ρ-plane n linearly independent solutions analytic in ρ, $u_1, \cdots, u_n$, expressible with all their derivatives in the form

$$u_i^{(k)}(x,\rho) = (\rho\omega_i)^k e^{\rho\omega_i x} \left\{ 1 + \sum_{l=1}^{l=m} \frac{A_{lk}(x)}{(\rho\omega_i)^l} + \frac{E_{ik}(x,\rho)}{\rho^{m+1}} \right\}$$
$$(i = 1, \cdots, n; k,m = 0,1,2, \cdots),$$

where $A_{lk}(x)$ is continuous together with its derivatives of all orders and $E_{ik}(x, \rho)$ is bounded for all ρ on T, $|\rho| \geqq R$, and all x on $(0, 1)$. The functions $A_{lk}(x)$ may be taken independent of the particular region T.

To demonstrate this theorem we first determine a function

$$\bar{u}_1(x,\rho) = e^{\rho\omega_1 x} \left\{ 1 + \sum_{l=1}^{l=m} \frac{A_{l0}(x)}{(\rho\omega_1)^l} \right\}$$

with the property that when substituted in the differential expression which is the left-hand member in our differential equation it gives rise to a series in powers of ρ beginning with a term in $(\rho\omega_1)^{n-m-2}$ at most. Since the substitution of $\rho\omega_i/\omega_1$ for ρ in the differential expression does not change it, the functions $\bar{u}_i=\bar{u}_1(x, \rho\omega_i/\omega_1)$ have the same property. Taking $\bar{u}_1, \cdots, \bar{u}_n$ as approximants to $u_1, \cdots, u_n$ we then apply the reasoning of Birkhoff's paper to establish the theorem for $k=0, \cdots, n-1$. For $k=n$ we substitute the asymptotic forms thus obtained in the differential equation; for $k=n+1$ we differentiate the equation once with respect to x and substitute the forms previously determined; continuing thus, we establish the theorem for all values of k by induction. Since the same functions $\bar{u}$, numbered in accordance with the ordering of the constants ω, can be employed for any other region T, the functions $A_{ik}(x)$ are independent of the particular region.

We may notice that, if we determine beforehand the integer n and the greatest value of k which we wish to employ, the functions $p_2, \cdots, p_n$ do not need to be taken as continuous with continuous derivatives of *all* orders in applications of this theorem. For the sake of simplifying a discussion sufficiently complicated in other ways we make this more restrictive assumption.

II. The general problem

The general problem which we shall discuss arises from a linear homogeneous differential system of the nth order:

$$u^{(n)} + \bigstar + p_2 u^{(n-2)} + \cdots + (p_n + \lambda)u = 0, \quad 0 \leqq x \leqq 1,$$
$$W_1(u) = 0, \qquad \cdots \qquad , W_n(u) = 0,$$

where $W_1(u), \cdots, W_n(u)$ are linearly independent homogeneous linear forms in $u(0), \cdots, u^{(n-1)}(0), u(1), \cdots, u^{(n-1)}(1)$ with real or complex constant coefficients. There is no loss of generality in restricting attention to the unit interval. As we have already pointed out in §I there is associated with this system a Green's function except when the boundary conditions are identically satisfied by some solution of the differential equation; this Green's function will have the essential properties of the Green's function defined in the case of an equation with continuous coefficients. We now make the assumption that the Green's function $G(x, y; \lambda)$ has infinitely many poles in the λ-plane, $\lambda_1, \lambda_2, \cdots$, which can be arranged so that $|\lambda_1| \leqq |\lambda_2| \leqq \cdots, \lim_{\nu\to\infty}|\lambda_\nu| = \infty$. We let $R_\nu(x, y)$ be the residue at $\lambda=\lambda_\nu$; considered as a function of x, this residue satisfies the differential system for $\lambda=\lambda_\nu$. We next define a system of circles $C_1, C_2, \cdots$, with centers at $\lambda=0$ and radii $\Lambda_1, \Lambda_2, \cdots$ forming a monotone sequence with limit $+\infty$. We may suppose

further that the region between two consecutive circles of the system contains at least one pole of G and as few others as possible. No pole shall lie on any of the circles C. For any function $f(x)$ summable on $(0, 1)$, and for that branch of the function $[1-(\lambda^4/\Lambda_\nu^4)]^{k+l}$ reducing to 1 at the origin, we consider the behavior of the integrals

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\,\lambda)d\lambda dy, \quad l \geqq 0,$$

$$\frac{1}{2\pi i}\int_0^1 f(y)\int_{C_\nu} G(y,x\,;\,\lambda)d\lambda dy,$$

as $\nu\to\infty$. By the theory of residues these expressions are respectively

$$\sum_{\alpha=1}^{\alpha=n_\nu}\left(1-\frac{\lambda_\alpha^4}{\Lambda_\nu^4}\right)^{k+l}\int_0^1 f(y)\frac{\partial^k}{\partial x^k}R_\alpha(x,y)dy,$$

$$\sum_{\alpha=1}^{\alpha=n_\nu}\int_0^1 f(y)R_\alpha(y,x)dy,$$

where n_ν is the number of poles of the Green's function in the circle C_ν.

To compare these expressions, suitably restricted in the case $k\neq 0$, with corresponding expressions in the theory of Fourier series we construct a differential system of order n for which the integrals and series just described become Fourier series. We then study the difference of the corresponding expressions formed for the two differential systems.

We shall, in the present paper, restrict ourselves to those differential systems for which the boundary conditions $W_1, \cdots, W_n$, after being reduced to their *normal* form, are *regular* according to Birkhoff's definition.* The differential system from which we obtain the Fourier series will be of this type. At a later time we shall come back to the *irregular cases*, $n=2$.

Our fundamental tool for this investigation is a theorem due to Lebesgue.† It is sufficient for our purposes to quote it, with minor modifications.

THEOREM IV. *Let there be given a function $\phi(x, y, \nu)$ defined for x and y on the interval (a, b) and for real positive values of ν belonging to a set N one of whose limit points is $+\infty$; and let ϕ be summable in y for each pair of values (x, ν). Then a sufficient condition that, as ν becomes infinite in any manner in N,*

$$\lim_{\nu\to\infty}\int_a^b f(y)\phi(x,y,\nu)dy = 0$$

* For the terms normal and regular, see Birkhoff, these Transactions, vol. 9 (1908), pp. 382–383.

† Lebesgue, Annales de Toulouse, ser. 3, vol. 1 (1909), pp. 52–55.

uniformly for all x belonging to a set E on (a, b), $f(x)$ being any summable function, is that

(1) $|\phi(x, y, \nu)| \leqq M$ *for all ν in N and for all x in E, except for values of y on a set E' of zero measure;*

(2) $\lim_{\nu\to\infty}\int_\alpha^\beta \phi(x, y, \nu)dy = 0$ *uniformly for x in E, $a \leqq \alpha < \beta \leqq b$.*

Lebesgue also shows these conditions necessary, but we shall not use this fact.

III. Fourier series

We find in the following theorem a generalisation of facts well known in the cases $n=1$ and $n=2$:

Theorem V. *The expansion problem associated with the differential system*

$$u^{(n)} + \lambda u = 0,$$
$$u(0) - u(1) = 0,$$
$$\cdots\cdots\cdots$$
$$u^{(n-1)}(0) - u^{(n-1)}(1) = 0,$$

in which the boundary conditions are regular, gives rise to Fourier series.

The regularity of the boundary conditions is a matter of direct computation which we omit. For the case where n is even this has been carried through elsewhere.*

We first determine the values of λ for which the differential system has solutions not identically zero. When $\lambda=0$, it is clear that the only solution of the system is the solution $u=$constant. When $\lambda\neq 0$ we make the substitution $\lambda=\rho^n$. Choosing n linearly independent solutions of the differential equation, $u_1=e^{\rho\omega_1 x}, \cdots, u_n=e^{\rho\omega_n x}$, we find all the other characteristic values of $\lambda=\rho^n$ from the roots of the equation

$$\begin{vmatrix} u_1(0) - u_1(1) & \cdots & u_n(0)-u_n(1) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ u_1^{(n-1)}(0) - u_1^{(n-1)}(1) & \cdots & u_n^{(n-1)}(0) - u_n^{(n-1)}(1) \end{vmatrix}$$
$$\equiv \Omega_n \rho^{n(n-1)/2}(1 - e^{\rho\omega_1}) \cdots (1 - e^{\rho\omega_n})$$

where Ω_n is a constant not zero. The roots of this equation other than $\rho=0$ are $\rho=2\nu\pi i/\omega_k$, $k=1, \cdots, n$, $\nu=\pm 1, \pm 2, \cdots$. The corresponding values of λ are $\lambda=-(2\nu\pi i)^n$, $\nu=\pm 1, \pm 2, \cdots$.

* Tamarkin, Rendiconti del Circolo Matematico di Palermo, vol. 34 (1912), p. 359.

Hence the Green's function $G(x, y; \lambda)$ exists for this differential system.

We consider separately the two cases $n=2\mu-1$, $n=2\mu$.

Case I. $n=2\mu-1$. Corresponding to $\lambda=\lambda_\nu=-(2\nu\pi i)^n$, $G(x, y, \lambda)$ has a residue $R_\nu(x, y)$; and the differential system a solution $e^{2\pi ix}$, $\nu=0, \pm 1, \cdots$. We use these facts to obtain a development of $R_\nu(x, y)$, considered as a function of y, in Fourier series. By a fundamental property of Green's function we have

$$(\lambda - \lambda_{\nu'}) \int_0^1 e^{2\nu'\pi iy} G(x, y; \lambda) dy = e^{2\nu'\pi ix}.$$

We divide this equation by $2\pi i(\lambda-\lambda_{\nu'})$ and integrate both sides of the resulting equality with respect to λ over a small circle surrounding the point $\lambda=\lambda_\nu$. We obtain

$$\int_0^1 e^{2\nu'\pi iy} R_\nu(x, y) dy = \begin{Bmatrix} e^{2\nu\pi ix}, & \nu' = \nu \\ 0 \quad , & \nu' \neq \nu \end{Bmatrix}.$$

If we form the corresponding equations for $-\nu'$, add them to those just written down and divide by two, we find

$$\int_0^1 \cos 2\nu'\pi y R_\nu(x, y) dy = \begin{Bmatrix} \frac{1}{2} e^{2\nu\pi ix}, & \nu' = \pm \nu \\ 0 \quad , & \nu' \neq \pm \nu \end{Bmatrix}.$$

In the same way

$$\int_0^1 \sin 2\nu'\pi y R_\nu(x, y) dy = \begin{Bmatrix} -\dfrac{i}{2} e^{2\nu\pi ix}, & \nu' = +\nu \\ +\dfrac{i}{2} e^{2\nu\pi ix}, & \nu' = -\nu \\ 0 \quad , & \nu' \neq \pm \nu \end{Bmatrix}.$$

Hence we see that $R_\nu(x, y) = e^{2\nu\pi i(x-y)}$, $\nu=0, \pm 1, \pm 2, \cdots$. If now we let C_0 be a circle with center at the origin in the λ-plane, including no other pole of $G(x, y; \lambda)$, C_1 a concentric circle including also λ_1, λ_{-1}, and so on, we have to consider the expressions

$$\frac{1}{2\pi i} \frac{\partial^k}{\partial x^k} \int_0^1 f(y) \int_{C_\nu} \left(1 - \frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x, y; \lambda) d\lambda dy,$$

$$\frac{1}{2\pi i} \int_0^1 f(y) \int_{C_\nu} G(y, x; \lambda) d\lambda dy.$$

On writing $\Lambda_\nu = (2\pi\omega_\nu)^n$ and denoting by $[\omega_\nu]$ the greatest integer less than ω_ν, we see that the first integral reduces to

$$\int_0^1 f(y)dy + 2\sum_{\alpha=1}^{\alpha=[\omega_\nu]}\left(1 - \frac{\alpha^{4n}}{\omega_\nu^{4n}}\right)^{k+l}\int_0^1 f(y)\frac{\partial^k}{\partial x^k}\cos 2\alpha\pi(x-y)dy,$$

and the second to a similar series. We recognize the Fourier series for the interval (0,1).

Case II. $n=2\mu$. For $\lambda=\lambda_\nu=-(2\nu\pi i)^n$, $\nu=0, 1, 2, \cdots$, the Green's function has residues $R_\nu(x, y)$; and the differential system has solutions $\cos 2\nu\pi x$, $\sin 2\nu\pi x$. From the identities

$$(\lambda - \lambda_{\nu'})\int_0^1 \cos 2\nu'\pi y G(x,y\,;\lambda)dy = \cos 2\nu'\pi y,$$

$$(\lambda - \lambda_{\nu'})\int_0^1 \sin 2\nu'\pi y G(x,y\,;\lambda)dy = \sin 2\nu'\pi x,$$

we obtain, as in Case I, the Fourier development of $R_\nu(x, y)$ considered as a function of y. It turns out that

$$R_0(x,y) = 1, \quad R_\nu(x,y) = 2\cos 2\nu\pi(x-y), \quad \nu = 1,2,\cdots.$$

As in Case I, the expansion problems lead us directly to Fourier series for the interval (0, 1).

To make clear the nature of the expressions involving factors of the form $[1-(\alpha^{4n}/\omega^{4n})]^{k+l}$ we give references to the literature from which the following theorem may be proved:

THEOREM VI. *Let $u_0(x)+u_1(x)+u_2(x)+\cdots$ be an infinite series whose terms are real functions of the real variable x, $a \leqq x \leqq b$. Then the expressions*

$$S_\Omega^{\delta,\Delta}(x) \equiv \sum_{\alpha=0}^{[\omega]}\left(1 - \frac{\alpha^\Delta}{\Omega}\right)^\delta u_\alpha(x)$$

where $\Omega=\omega^\Delta$, $\Delta>0$, $\delta\geqq 0$, are the Riesz typical means of the first kind of order δ and type α^Δ for the infinite series in question. If $\lim_{\Omega\to\infty} S_\Omega^{\delta,\Delta}$ exists and is equal to $U(x)$, the infinite series is said to be summable (α^Δ,δ) at the point x to the value $U(x)$. Summability (α^Δ, δ) implies summability $(\alpha^{\Delta'}, \delta)$ to the same value; if the summability (α^Δ, δ) is uniform for a certain range of values of x, then the summability $(\alpha^{\Delta'}, \delta)$ is also uniform for that range. Likewise, summability (α^Δ, δ) implies summability (α^Δ, δ') to the same value if $\delta' \geqq \delta$, with similar remarks on uniform summability. The means $S_\Omega^{\delta,1}(x)$ are equivalent to the Cesàro means for integral δ and their generalisations for arbitrary δ; that is to say, summability (α, δ) implies summability $C(\delta)$ to the same value, and conversely, with remarks on uniform summability like those previously made.

Riesz's typical means are discussed in some detail in a joint work of Hardy and Riesz,* where further references are given.† The series there considered are series of constants; but the extension of the methods employed to questions of uniform summability when series of functions are involved requires merely appropriate modification of Lemmas 4, 5, 7, 8. Theorem 16 gives the relation between summability $(\alpha^{\Delta}, \delta)$ and summability $(\alpha^{\Delta}, \delta')$ when $\delta' \geqq \delta$. The relation between summability $(\alpha^{\Delta}, \delta)$ and summability $(\alpha^{\Delta'}, \delta)$ is a special case of the theorem stated in the second foot-note on page 33; a proof of this fact could be modeled on that of Theorem 17. The comparison of summability (α, δ) with summability $C(\delta)$ follows a paper of Riesz.‡ We add the

Corollary I. *Summability $(\alpha^{\Delta}, \delta)$ and summability $C(\delta)$ are completely equivalent with respect to summability at a point and uniform summability over a set of points; each implies the other.*

The importance of this corollary is due to the fact that the known results concerning the Cesàro summability of Fourier series and its term-by-term derived series can be carried over to summability by the Riesz means, which are better suited to the form of Fourier series which we shall discuss under Theorem V. We shall then be able to carry these results over to the general Birkhoff series by the intervention of this corollary and Theorem V.

IV. Birkhoff series for $n=2\mu-1$

Distinct differences between the case when n is odd and the case when n is even make it necessary to consider each case separately. The methods employed in the two cases are virtually the same; we therefore give a complete study of the case $n=2\mu-1$ and then describe the necessary changes in mode of attack and in statement of theorems in the case $n=2\mu$.

By our hypotheses, the differential system considered is

$$u^{(n)} + \bigstar + p_2 u^{(n-2)} + \cdots + (p_n + \lambda)u = 0,$$
$$W_1(u) = 0, \cdots, W_n(u) = 0, \; n = 2\mu - 1,$$

---

* Hardy and Marcel Riesz, *The General Theory of Dirichlet's Series* (Cambridge Tracts, No. 18, 1915), Chapters IV and V.

† See especially Marcel Riesz, *Sur les séries de Dirichlet et les séries entières*, Comptes Rendus, vol. 149 (1909), pp. 909–912.

‡ Marcel Riesz, *Une méthode de sommation équivalente à la méthode des moyens arithmétiques*, Comptes Rendus, vol. 152 (1911), pp. 1651–1654.

where

$$W_i(u) \equiv \alpha_i u^{(k_i)}(0) + \beta_i u^{(k_i)}(1) + \cdots,$$

$$n - 1 \geqq k_1 \geqq k_2 \geqq \cdots \geqq k_n, \quad k_i \neq k_{i+2},$$

and $\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_n, k_1, \cdots, k_n$ are subjected to certain conditions in accordance with the definition of regularity.* On making the substitution $\lambda = \rho^n$ we are led to consider instead of the entire λ-plane a corresponding sector of the ρ-plane, which we shall take as composed of two adjacent sectors S: $\mathcal{R}(\rho\omega_1) \leqq \mathcal{R}(\rho\omega_2) \leqq \cdots \leqq \mathcal{R}(\rho\omega_n)$, arising, of course, from distinct orderings of the constants ω; this region we call Σ. The sector S is a region T of §I for $c=0$. We recall now the facts concerning the asymptotic distribution of the values of ρ giving rise to the characteristic values of λ. On an arbitrary sector S these values of ρ lie asymptotically near points equally spaced at distance 2π along a line parallel to the bisecting ray of S; there is one and only one value of ρ near each of these points when $|\rho|$ is large; and the corresponding values of λ are simple characteristic values of the differential system.†

We wish to compare any two given differential systems of order $n = 2\mu - 1$. From the ρ-plane we remove all the points which give characteristic values of one system or the other of the two which we are comparing; we do this by removing the interiors of circles σ of arbitrarily small radius ϵ, one described about each such point as center. The portions of S, Σ which remain we denote by S', Σ'. We then denote by Γ any circular arc on Σ' with center at the origin terminated by the rectilinear portions of the boundary of Σ'. The image of Γ in the λ-plane is a circle C; the totality of such circles C forms an infinite set of concentric annular regions, none of which contains a characteristic value. From the behavior of the characteristic values for large $|\lambda|$ we see that the large circles of the set $C_1, C_2, C_3, \cdots$ as described in §II can be selected as circles C; this is true simultaneously for the two differential systems we are considering.

The region S is such that on it‡

$$\mathcal{R}(\rho\omega_1) \leqq \mathcal{R}(\rho\omega_2) \leqq \cdots \leqq \mathcal{R}(\rho\omega_{\mu-1}) \leqq -\beta < 0;$$

$$\mathcal{R}(\rho\omega_\mu) = 0 \text{ on the bisecting ray of } S;$$

$$\mathcal{R}(\rho\omega_n) \geqq \cdots \geqq \mathcal{R}(\rho\omega_{\mu+1}) \geqq \beta > 0.$$

* Birkhoff, these Transactions, vol. 9 (1908), p. 383.
† Birkhoff, these Transactions, vol. 9 (1908), pp. 383–386.
‡ Birkhoff, these Transactions, vol. 9 (1908), p. 384.

We therefore denote by S_1 that part of S for which $\Re(\rho\omega_\mu) \leqq 0$; the part for which $\Re(\rho\omega_\mu) \geqq 0$ by S_2. The notations S_1', S_2' can be introduced according to the principle previously employed. The portions of Γ on S', S_1', S_2', we shall call γ, γ_1, γ_2 respectively. The common radius of these arcs we denote by R.

We now prove several preparatory lemmas.

LEMMA III. *If $m(x, y, \rho)$ is a function of x, y, and ρ bounded for $a \leqq x \leqq b$, $a \leqq y \leqq b$, and for all ρ on S' such that $\theta' \leqq \arg \rho \leqq \theta''$; and if α is any arc of γ between θ' and θ'', then the integral $\int_\alpha m(x, y, \rho)\, d\rho/\rho$ is a bounded function of x, $a \leqq x \leqq b$, y, $a \leqq y \leqq b$, and R.*

We have

$$\left| \int_\alpha m \frac{d\rho}{\rho} \right| \leqq M \int_\alpha d\theta \leqq M(\theta'' - \theta').$$

LEMMA IV. *If $m(x, \rho)$ is a function of x and ρ bounded for $a \leqq x \leqq b$, $0 < a < b < 1$, and for all ρ on S' such that $\theta' \leqq \arg \rho \leqq \theta''$, then*

$$\lim_{R\to\infty} \int_\alpha \rho^k e^{\rho\omega_i x} m(x,\rho) d\rho = 0 \qquad (i = 1, \cdots, \mu - 1),$$

$$\lim_{R\to\infty} \int_\alpha \rho^k e^{\rho\omega_i (x-1)} m(x,\rho) d\rho = 0 \qquad (i = \mu + 1, \cdots, n),$$

uniformly, $0 < a \leqq x \leqq b < 1$, for any positive k.

We let δ be the lesser of a, $1-b$. Then

$$\left| \int_\alpha \rho^k e^{\rho\omega_i x} m d\rho \right| \leqq M R^{k+1} e^{-R\beta\delta} \int_\alpha d\theta \leqq M R^{k+1} e^{-R\beta\delta}(\theta'' - \theta') \to 0$$

when $i = 1, \cdots, \mu - 1$; and

$$\left| \int_\alpha \rho^k e^{\rho\omega_i (x-1)} m d\rho \right| \leqq M(\theta'' - \theta') R^{k+1} e^{-R\beta\delta} \to 0$$

when $i = \mu + 1, \cdots, n$.

LEMMA V. *If $m(x, \rho)$ is a function of x and ρ bounded for $0 < a \leqq x \leqq b < 1$, and for all ρ on S_1', then*

$$\int_\alpha \rho^k \left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^{k+l} e^{\rho\omega_\mu x} m(x,\rho) d\rho$$

for any $k \geqq 0$, (1) is uniformly bounded for all α on γ_1 and all x on (a, b) if $l \geqq 0$, and (2) approaches zero uniformly as $R \to \infty$ if $l > 0$. If $m(x,\rho)$ is a function

of x and ρ bounded for x on (a, b) and for all ρ on S_2' similar statements are true of the integral

$$\int_\alpha \rho^k\left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^{k+l} e^{\rho\omega_\mu(x-1)}m(x,\rho)d\rho.$$

In the case of the first integral we introduce an angle ϕ measured from the bisecting ray of S as initial direction, with the positive sense in S_1. We then see that we can write

$$\rho\omega_\mu = iRe^{i\phi},$$
$$0 \leqq \phi/2 \leqq \sin\phi, \quad 0 \leqq \phi \leqq \frac{\pi}{2n},$$
$$0 \leqq |\, 1 - e^{4ni\phi}\,|^{k+l} \leqq N\phi^{k+l}, \quad 0 \leqq \phi \leqq \frac{\pi}{2n}\cdot$$

We find

$$\left|\int_\alpha \rho^k\left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^{k+l} e^{\rho\omega_\mu x}m d\rho\right|$$
$$= \left|\int_{\phi'}^{\phi''} R^{k+1}(1 - e^{4ni\phi})^{k+l}e^{(k+1)i\phi}e^{-Rx\sin\phi}e^{iRx\cos\phi}m d\phi\right|$$
$$\leqq M\int_0^{\pi/2n} R^{k+1}\,|\,1 - e^{4ni\phi}\,|^{k+l}e^{-R\delta\phi/2}d\phi \leqq MN\int_0^{\pi/2n} R^{k+1}\phi^{k+l}e^{-R\delta\phi/2}d\phi$$
$$= \frac{MN}{R^l}\int_0^{\pi R/2n} \xi^{k+l}e^{-\delta\xi/2}d\xi \leqq \frac{MN}{R^l}\int_0^\infty \xi^{k+l}e^{-\delta\xi/2}d\xi = \frac{K}{R^l}\cdot$$

This establishes the first result. The second integral can be treated in precisely similar fashion.

Lemma VI. *If $m(x, y, \rho)$ is bounded $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$, for all ρ on S, then*

$$\int_\alpha^x \int_{\gamma_1} e^{\rho\omega_i(x-y)}\,\frac{m(x,y,\rho)}{\rho}\,d\rho dy$$

approaches zero uniformly, $i = 1, \cdots, \mu$, $0 \leqq \alpha \leqq x \leqq 1$, and $i = \mu+1, \cdots, n$, $0 \leqq x \leqq \alpha \leqq 1$; and

$$\int_\alpha^x \int_{\gamma_2} e^{\rho\omega_i(x-y)}\,\frac{m(x,y,\rho)}{\rho}\,d\rho dy$$

approaches zero uniformly, $i = 1, \cdots, \mu-1$, $0 \leqq \alpha \leqq x \leqq 1$, and $i = \mu, \cdots, n$, $0 \leqq x \leqq \alpha \leqq 1$.

We have when $i=1,\cdots,\mu-1$, $0\leqq\alpha\leqq x\leqq 1$,

$$\left|\int_\alpha^x\int_{\gamma_1} e^{\rho\omega_i(x-y)}\frac{m}{\rho}\,d\rho dy\right|\leqq\int_\alpha^x\int_0^{\pi/2n} e^{-R\beta(x-y)}M\,d\phi dy$$

$$=\frac{\pi M}{2n}\,\frac{1-e^{-R\beta(x-\alpha)}}{R\beta}\leqq\frac{\pi M}{2R\beta n}\to 0.$$

Similar reasoning applies when $i=\mu+1,\cdots,n$, $0\leqq x\leqq\alpha\leqq 1$. The integrals over γ_2, $i\neq\mu$, are discussed in the same way.

When $i=\mu$ the treatment is slightly different. For $0\leqq\alpha\leqq x\leqq 1$

$$\left|\int_\alpha^x\int_{\gamma_1} e^{\rho\omega_\mu(x-y)}\frac{m}{\rho}\,d\rho dy\right|\leqq\int_\alpha^x\int_0^{\pi/2n} e^{-R\phi(x-y)/2}M\,d\phi dy$$

$$=M\int_\alpha^x\frac{1-e^{-\pi R(x-y)/4n}}{R(x-y)}\,dy=\frac{M}{R}\int_\alpha^{R(x-\alpha)}\frac{1-e^{-\pi\xi/4n}}{\xi}\,d\xi$$

$$\leqq\frac{M}{R}\int_0^R\frac{1-e^{-\pi\xi/4n}}{\xi}\,d\xi=\frac{M}{R}O(\log R)\to 0.$$

By analogous means we obtain the corresponding result for the integral over γ_2.

We can now demonstrate

THEOREM VII. *If G and $\bar G$ are the Green's functions for any two regular differential systems of order $n=2\mu-1$, then for any region S' the integral*

$$\int_\gamma n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l\{G(x,y;\rho^n)-\bar G(x,y;\rho^n)\}\,d\rho$$

is bounded for all γ on S', $0\leqq y\leqq 1$, $0<a\leqq x\leqq b<1$, for $l\geqq 0$.

We first recall the explicit form of the Green's function, which we know to exist in the present case. We have

$$G(x,y;\rho^n)\equiv\frac{(-1)^n\begin{vmatrix}u_1(x)&\cdots&u_n(x)&g(x,y)\\W_1(u_1)&\cdots&W_1(u_n)&W_1(g)\\\cdots&\cdots&\cdots&\cdots\\W_n(u_1)&\cdots&W_n(u_n)&W_n(g)\end{vmatrix}}{\begin{vmatrix}W_1(u_1)&\cdots&W_1(u_n)\\\cdots&\cdots&\cdots\\W_n(u_1)&\cdots&W_n(u_n)\end{vmatrix}}$$

where $u_1, \cdots, u_n$ are n linearly independent solutions of the differential equation and

$$g(x,y) \equiv \left\{\frac{1}{2}\sum_{i=1}^{i=n} u_i(x)v_i(y)\ ;\ -\frac{1}{2}\sum_{i=1}^{i=n} u_i(x)v_i(y)\right\},$$

the notation $\{A; B\}$ indicating that A is to be taken if $x \geqq y$, B if $x \leqq y$, and the functions $v_1, \cdots, v_n$ being defined by the identity in $\tau_1, \cdots, \tau_n$

$$\tau_1 v_1(y) + \cdots + \tau_n v_n(y) \equiv \frac{\begin{vmatrix} u_1(y) & \cdots & u_n(y) \\ \cdots & \cdots & \cdots \\ u_1^{(n-2)}(y) & \cdots & u_n^{(n-2)}(y) \\ \tau_1 & \cdots & \tau_n \end{vmatrix}}{\begin{vmatrix} u_1(y) & \cdots & u_n(y) \\ \cdots & \cdots & \cdots \\ u_1^{(n-1)}(y) & \cdots & u_n^{(n-1)}(y) \end{vmatrix}}.$$

We shall take $u_1, \cdots, u_n$ as the functions defined in Theorem III for the sector S. It is then possible to determine the asymptotic form of the Green's function on S_1' and S_2' in sufficient detail that the present theorem is apparent from the lemmas just established.

Case I. ρ on S_1'. We multiply the first μ columns in the numerator of G by $\frac{1}{2}v_1(y), \cdots, \frac{1}{2}v_\mu(y)$, respectively, the next $\mu-1$ columns by $-\frac{1}{2}v_{\mu+1}(y), \cdots, -\frac{1}{2}v_n(y)$ and add to the last. We then take the first term in the new last column outside the determinant. The result of substituting the asymp totic forms of $u_1, \cdots, u_n, v_1, \cdots, v_n$ in the expression thus obtained is*

$$n\rho^{n-1}G = \left\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}[\omega_i]\ ;\ +\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}[\omega_i]\right\} + \frac{\Delta_1}{[\theta_0]+e^{\rho\omega_i}[\theta_1]}$$

where Δ_1 is a determinant of order $n+1$, in which the element a_{hj}, $h=0, \cdots, n, j=1, \cdots, n+1$, may be described as follows:

$$a_{0j} = e^{\rho\omega_j x}[1], j = 1, \cdots, \mu\ ;\ a_{0,j} = e^{\rho\omega_j(x-1)}[1], j = \mu+1, \cdots, n\ ;$$
$$a_{0,n+1} = 0\ ;$$

and for $h=1, \cdots, n$

$$a_{hj} = [\alpha_h\omega_j^{k_h}], j = 1, \cdots, \mu-1\ ;\ a_{h\mu} = [\alpha_h\omega_\mu^{k_h}] + e^{\rho\omega_\mu}[\beta_h\omega_\mu^{k_h}]\ ;$$

$$a_{hj} = [\beta_h\omega_j^{k_h}], j = \mu+1, \cdots, n\ ;\ a_{h,n+1} = \sum_{i=1}^{i=\mu} e^{\rho\omega_i(1-y)}[\beta_h\omega_i^{k_h+1}]$$
$$- \sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}[\alpha_h\omega_i^{k_h+1}].$$

* Birkhoff, these Transactions, vol. 9 (1908) pp. 389–395.

The notation [ ] is used to indicate the leading term of the asymptotic form in question. The exponential terms in the determinant Δ_1 are bounded for all x and y on $(0, 1)$ and for all ρ on S_1; and $1/\{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]\}$, in which $\theta_0\,\theta_1\neq 0$ by the regularity of the boundary conditions, is bounded for all ρ on S_1'.* Hence we are able to write

$$\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l\left(n\rho^{n-1}G-\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\omega_i\,;+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\omega_i\right\}\right)$$

$$=\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\,;+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\right\}$$

$$+\sum_{i=1}^{i=\mu}e^{\rho\omega_i x}M_i(x,y,\rho)+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-1)}M_i(x,y,\rho)$$

where m and M denote functions bounded for all x and y on $(0, 1)$ and for all ρ on S_1'. Thus, by Lemmas III, IV, V we see that

$$\int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l\left(n\rho^{n-1}G-\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\omega_i\,;+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\omega_i\right\}d\rho\right.$$

is bounded $0\leqq y\leqq 1$, $0<a\leqq x\leqq b<1$, for all γ_1 on S_1'. The application of Lemma III, it should be observed, is made to the bracket term on the right, which is clearly of the form m/ρ for all x and y on $(0, 1)$ and all ρ on S_1.

Of course, in the result just obtained we may replace G by $\bar{G}$; a combina-t ionof the facts thus established shows that the integral

$$\int_{\gamma_1}n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l(G-\bar{G})d\rho$$

is bounded $0\leqq y\leqq 1$, $0<a\leqq x\leqq b<1$, for all γ_1 on S_1', $l\geqq 0$.

Case II. ρ on S_2'. We now multiply the first $\mu-1$ columns in the numerator determinant of G by $\frac{1}{2}v_1(y),\cdots,\frac{1}{2}v_{\mu-1}(y)$ respectively, the next μ columns by $-\frac{1}{2}v_\mu(y),\cdots,-\frac{1}{2}v_n(y)$, and add to the last. We then continue as in Case I, finding

$$n\rho^{n-1}G=\left\{-\sum_{i=1}^{i=\mu-1}e^{\rho\omega_i(x-y)}[\omega_i]\,;+\sum_{i=\mu}^{i=n}e^{\rho\omega_i(x-y)}[\omega_i]\right\}+\frac{\Delta_2}{[\theta_0]e^{-\rho\omega_\mu}+[\theta_1]}$$

where Δ_2 differs from Δ_1 in Case I in having the definition of $a_{h\mu}$ replaced by

$$a_{0,\mu}=e^{\rho\omega_\mu(x-1)}[1]\,;\ a_{h\mu}=e^{-\rho\omega_\mu}[\alpha_h\omega_\mu{}^{k_h}]+[\beta_h\omega_\mu{}^{k_h}],\ h=1,\cdots,n,$$

* Birkhoff, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), pp. 120–121.

and in having the sums in the definition of $a_{h,n+1}$, $h>0$, extended from $i=1$ to $i=\mu-1$ and from $i=\mu$ to $i=n$ respectively. The exponential terms in the determinant Δ_2 are bounded for all x and y on $(0, 1)$ and for all ρ on S_2'; and $1/\{[\theta_0]e^{-\rho\omega_\mu}+[\theta_1]\}$ is bounded for all ρ on S_1'. The reasoning now proceeds as in Case I. Thus we see that the integral

$$\int_{\gamma_2} n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l (G-\overline{G})d\rho$$

is bounded $0\leqq y\leqq 1$, $0<a\leqq x\leqq b<1$, for all γ_2 on S_2', $l\geqq 0$.

By combining the facts in Cases I and II we obtain the theorem as stated for the integral over γ on S'. Since S' is arbitrary the result holds also for the arc Γ on Σ' and can be interpreted in terms of the circles C in the λ-plane.

We next demonstrate

THEOREM VIII. *Under the hypotheses of Theorem* VII, *the integral*

$$\int_\alpha^\beta \int_\gamma n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l (G(x,y\,;\,\rho^n)-\overline{G}(x,y\,;\,\rho^n))d\rho dy,$$

where $l\geqq 0$ *and* $0\leqq\alpha<\beta\leqq 1$, *has the limit zero as* $R\to\infty$, *uniformly* $0<a\leqq x\leqq b<1$.

We split the integral over (α, β) into the sum of the integrals over (α, x) and (x, β); since the last two integrals are of the same form we consider but one of them.

Case I. ρ on γ_1. We employ the asymptotic form of $n\rho^{n-1}G$ obtained in the preceding theorem. In the determinant Δ_1 the variable y occurs only in the last column so that the integration with respect to y can be performed directly in the determinant. We see at once that

$$\int_x^\alpha e^{\rho\omega_i y}\left(1+\frac{m}{\rho}\right)dy = \frac{e^{\rho\omega_i(1-x)}-e^{\rho\omega_i(1-\alpha)}}{\rho\omega_i} + \int_x^\alpha e^{\rho\omega_i(1-y)}\frac{m}{\rho}dy = \frac{m(x,\alpha,\rho)}{R}$$

for all x and α on $(0, 1)$ and for ρ on γ_1, $i=1,\cdots,\mu$; and that

$$\int_x^\alpha e^{-\rho\omega_i y}\left(1+\frac{m}{\rho}\right)dy = \frac{m(x,\alpha,\rho)}{R}$$

for all x and α on $(0, 1)$ and for ρ on γ_1, $i=\mu+1,\cdots,n$. Hence we see that

$$\int_x^\alpha \int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l\left(n\rho^{n-1}G-\left\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}\omega_i\ ;\ +\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}\omega_i\right\}\right)d\rho dy$$

$$=\int_x^\alpha \int_{\gamma_1}\left\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\ ;+\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\right\}d\rho dy$$

$$+\frac{1}{R}\int_{\gamma_1}\left(\sum_{i=1}^{i=\mu} e^{\rho\omega_i x}m_{n+i}(x,\alpha,\rho)+\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-1)}m_{n+i}(x,\alpha,\rho)\right)d\rho.$$

The two expressions on the right of this equation approach zero uniformly as $R\to\infty$, $0<a\leqq x\leqq b<1$, by Lemmas IV, V, VI.

We can replace G by $\bar{G}$ in this result; and we may write β instead of α. Thus it follows that

$$\lim_{R\to\infty}\int_\alpha^\beta \int_{\gamma_1} n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l(G(x,y\ ;\ \rho^n)-\bar{G}(x,y\ ;\ \rho^n))d\rho dy=0$$

uniformly $0<a\leqq x\leqq b<1$.

Case II. ρ on γ_2. We must use here the asymptotic form of $n\rho^{n-1}G$ employed in Case II of the preceding theorem. We separate the integral over (α, β) into two parts as in Case I, and then perform the integration with respect to y in the last column of the determinant Δ_2. On expanding the determinant we have

$$\int_x^\alpha \int_{\gamma_2}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l\left(n\rho^{n-1}G-\left\{-\sum_{i=1}^{i=\mu-1} e^{\rho\omega_i(x-y)}\omega_i\ ;+\sum_{i=\mu}^{i=n} e^{\rho\omega_i(x-y)}\omega_i\right\}\right)d\rho dy$$

$$=\int_x^\alpha \int_{\gamma_2}\left\{-\sum_{i=1}^{i=\mu-1} e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\ ;+\sum_{i=\mu}^{i=n} e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\right\}d\rho dy$$

$$+\frac{1}{R}\int_{\gamma_2}\left(\sum_{i=1}^{i=\mu-1} e^{\rho\omega_i x}m_{n+i}(x,\alpha,\rho)+\sum_{i=\mu}^{i=n} e^{\rho\omega_i(x-1)}m_{n+i}(x,\alpha,\rho)\right)d\rho.$$

By Lemmas IV, V, VI, the expressions on the right approach zero uniformly as $R\to\infty$, $0<a\leqq x\leqq b<1$. Thus, as in Case I,

$$\lim_{R\to\infty}\int_\alpha^\beta \int_{\gamma_2} n\rho^{n-1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^l(G-\bar{G})d\rho dy=0$$

uniformly, $0<a\leqq x\leqq b<1$.

From Cases I and II we conclude the truth of the present theorem. By the arbitrariness of the sector S' the theorem remains true if the integration with respect to ρ takes place over Γ.

From the two theorems just established there follows

THEOREM IX. *If $f(x)$ is any summable function defined on* (0, 1) *and if $l \geqq 0$, then*

$$\lim_{R\to\infty} \int_0^1 f(y) \int_\gamma n\rho^{n-1}\left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^l (G(x,y;\rho^n) - \bar{G}(x,y;\rho^n))d\rho dy = 0,$$

$$\lim_{\Lambda\to\infty} \int_0^1 f(y) \int_C \left(1 - \frac{\lambda^4}{\Lambda^4}\right)^l (G(x,y;\lambda) - \bar{G}(x,y;\lambda))d\lambda dy = 0,$$

uniformly, $0 < a \leqq x \leqq b < 1$.

We see at once that Theorems VII and VIII together give us the sufficient conditions of Theorem IV and therefore establish the present theorem with regard to the first integral. Since the sector S' is arbitrary the arc γ may be replaced by Γ in the first integral; and then a return to the variable λ gives the second integral.

Theorem IX clearly yields a large amount of information concerning the behavior of the expansions arising from two distinct regular differential systems of order $n=2\mu-1$. There is one point, however, which we desire to examine before stating our final theorems on the comparison of these expansions. If we let N_Λ, $\bar{N}_\Lambda$ be the number of residues of the two Green's functions G and $\bar{G}$ respectively arising from poles in the circle C of radius Λ, we clearly have $|N_\Lambda - \bar{N}_\Lambda| \leqq K$ independent of Λ, though it is not necessary that $N_\Lambda = \bar{N}_\Lambda$. Thus the expansions determined by the integrals over C are not compared term-by-term in Theorem IX in general. In order to compare the sums given by the first N residues in the two cases we clearly need to obtain a generalization of the theorem of Riemann-Lebesgue, familiar in the theory of Fourier series: we need to show that the sum of any limited number of terms of the expansion becomes arbitrarily small when each term of the sum arises from a pole of G sufficiently removed from the origin. This is the aim of the theorems which succeed.

THEOREM X. *The residues of the Green's function for a regular differential system of order $n=2\mu-1$ are a set of functions of x and y uniformly bounded for all x and y on* (0, 1).

On the region S', $n\rho^{n-1}G(x, y; \rho^n)$ is uniformly bounded for all x and y on (0, 1), as we see from the asymptotic forms discussed in Theorem VII. Now we can compute any desired residue of the Green's function by evaluating the integral

$$R_\nu(x,y) = \frac{1}{2\pi i}\int_\sigma n\rho^{n-1}G(x,y\,;\,\rho^n)d\rho$$

taken over the circumference of the circle σ of radius ϵ described about the corresponding pole of G in the region Σ; the circles σ remote from the origin lie entirely interior to a region S. Thus we find $|R_\nu(x,\,y)| \leqq M\epsilon$ for all x and y on $(0,\,1)$, $\nu = 1,\,2,\,\cdots$.

THEOREM XI. *Under the hypotheses of Theorem* X, *the integral*

$$\frac{1}{2\pi i}\int_\alpha^\beta\int_\sigma n\rho^{n-1}G(x,y\,;\,\rho^n)d\rho dy = \int_\alpha^\beta R_\nu(x,y)dy$$

approaches zero uniformly, $0 \leqq x \leqq 1$, *as the center of the circle* σ *recedes indefinitely from the origin. We suppose* $0 \leqq \alpha < \beta \leqq 1$.

As in Theorem VIII we divide the integral over $(\alpha,\,\beta)$ into integrals over $(\alpha,\,x)$ and $(x,\,\beta)$, of which we consider only one in detail. We then show that on S' we have

$$\int_x^\alpha n\rho^{n-1}G(x,y\,;\,\rho^n)dy = \frac{m(x,\alpha,\rho)}{\rho}\,.$$

It follows that

$$\left|\int_\sigma\int_\alpha^\beta n\rho^{n-1}G(x,y\,;\,\rho^n)dy\right| \leqq \frac{2\pi M\epsilon}{R'}$$

where R' is the shortest distance from the origin to the circle σ.

Case I. ρ on S_1'. We refer once again to the asymptotic form of $n\rho^{n-1}G(x,\,y;\,\rho^n)$ given in Case I of Theorem VII. We perform the integration with respect to y directly in the determinant Δ_1 as in Theorem VIII and find that the contribution of this term is of the form $m(x,\,\alpha,\,\rho)/\rho$, just as before. It remains to treat the bracket terms in the asymptotic form. We have

$$\int_x^\alpha e^{\rho\omega_i(x-y)}\left(\omega_i + \frac{m_i}{\rho}\right)dy = \frac{1 - e^{\rho\omega_i(x-\alpha)}}{\rho} + \int_x^\alpha e^{\rho\omega_i(x-y)}\frac{m_i}{\rho}\,dy = \frac{m_i(x,\alpha,\rho)}{\rho}$$

for ρ on S_1', $0 \leqq \alpha \leqq x \leqq 1$, $i = 1,\,\cdots,\,\mu$; and

$$\int_x^\alpha e^{\rho\omega_i(x-y)}\left(\omega_i + \frac{m_i}{\rho}\right)dy = \frac{m_i(x,\alpha,\rho)}{\rho}$$

for ρ on S_1', $0 \leqq x \leqq \alpha \leqq 1$, $i = \mu+1,\,\cdots,\,n$. Hence, the bracket terms also contribute an expression of the form $m(x,\,\alpha,\,\rho)/\rho$, ρ on S_1', for all x and α on $(0,\,1)$. Consequently on S_1' we have

$$\int_x^\alpha n\rho^{n-1}G(x,y\,;\,\rho^n)dy = \frac{m(x,\alpha,\rho)}{\rho}\,.$$

Case II. ρ on S_2'. The procedure is the same as in Case I except for the difference of the two sectors with regard to the constant ω_μ.

As we have already pointed out, the theorem flows at once from the results of Cases I and II.

We now obtain the generalised theorem of Riemann-Lebesgue:

THEOREM XII. *If the hypotheses of Theorem* X *are fulfilled, and if* $f(x)$ *is any summable function on* (0, 1), *then*

$$\frac{1}{2\pi i}\int_0^1 f(y)\int_\sigma n\rho^{n-1}G(x,y\,;\,\rho^n)d\rho dy$$

approaches zero uniformly, $0\leqq x\leqq 1$, *as the center of the circle* σ *recedes indefinitely from the origin. In other words, if* $R_\nu(x, y)$ *is the residue of* $G(x, y;\lambda)$ *corresponding to the pole* $\lambda=\lambda_\nu$, *then*

$$\lim_{\nu\to\infty}\int_0^1 f(y)R_\nu(x,y)dy = 0$$

uniformly, $0\leqq x\leqq 1$.

The theorem is established at once by comparing Theorems IV, X, XI.

We can now make the following fundamental assertion:*

THEOREM XIII. *On any closed interval* (a, b) *completely interior to* (0, 1) *the term-by-term difference of the expansions formed for any summable function* $f(x)$ *in connection with two regular differential systems of order* $n=2\mu-1$ *converges uniformly to zero. These series we shall call Birkhoff series of order* $n=2\mu-1$; *and we may restate the preceding result, saying that any two Birkhoff series of order* $n=2\mu-1$ *are equivalent on* (a, b). *In particular, Birkhoff series of order* $n=2\mu-1$ *are equivalent to Fourier series on* (a, b).

The material on which the proof rests is to be found in Theorems V, IX, XII. If we take $l=0$ in Theorem IX we find that

$$\frac{1}{2\pi i}\int_0^1 f(y)\int_C (G(x,y\,;\,\lambda) - \bar{G}(x,y\,;\,\lambda))d\lambda dy$$

gives the difference of N_Λ terms of one series and $\bar{N}_\Lambda$ of the other, and approaches zero uniformly on (a, b) as $\Lambda\to\infty$. By Theorem XII we can re-

* Tamarkin, D. E., Chapter V, has shown the truth of the last statement here, without actually discussing the term-by-term difference of the two series; his statement of the result is found in Lemma 6, where we have $p=1$, $\omega_k(x)=1/n$, $q=2\pi$, $|c_k|=1$.

move a limited number of terms from each series without affecting the uniform convergence of their difference to zero, provided that the terms removed correspond to poles of the Green's functions which recede indefinitely from the origin. Because of the regular distribution of the poles of the two Green's functions, we can always do this in such wise that we obtain the term-by-term difference of the first N terms of the two series, which therefore converges uniformly to zero on (a, b) as N becomes infinite, passing through all positive integral values. Theorem V tells us that the Fourier series is a special case of the Birkhoff series of order $n=2\mu-1$, and therefore justifies the last statement of the theorem.

The consequences of this theorem are so important that it is worthwhile to describe them in detail. The theorem means essentially that on any interval (a, b) interior to $(0, 1)$ any two Birkhoff series of order $n=2\mu-1$ have exactly the same behavior and both behave like Fourier series. In particular, properties of Fourier series such as uniform and non-uniform convergence, divergence, oscillation, uniform and non-uniform summability, are carried over directly to Birkhoff series. The various tests for the convergence and summability of Fourier series apply also to Birkhoff series. We conclude the existence of continuous functions whose Birkhoff series converge non-uniformly or diverge. Gibbs' phenomenon is common to Fourier series and Birkhoff series. The theorems of Fejér-Chapman and Lebesgue-Hardy are true for Birkhoff series. Another important fact which we may observe is that the behavior of Birkhoff series at any interior point of $(0, 1)$ depends only upon the nature of the expanded function in the neighborhood of that point.

It is also in place to note that, for the class of all summable functions, this result is the strongest that can be obtained. The equivalence cannot be extended to the entire interval $(0, 1)$ as can be shown by an example.*

In the preceding paragraphs we mentioned the theorems of Fejér-Chapman and Lebesgue-Hardy, which are concerned with the Cesàro summability of Fourier series. It is not without interest to apply our methods to prove these theorems. To accomplish this we need a lemma whose truth is suggested by facts from the theory of Bessel functions.

LEMMA VII. *The function*

$$\Phi(\alpha) = \int_{-1}^{+1} (1 - \phi^4)^l e^{-i\alpha\phi} d\phi = 2\int_0^1 (1 - \phi^4)^l \cos \alpha\phi d\phi, \quad l \geqq 0,$$

---

* Stone, these Transactions, vol. 26 (1924), pp. 335–355, § II.

satisfies the inequalities

$$|\Phi(\alpha)| \leqq K, \quad 0 \leqq \alpha \leqq 1\,; \quad |\Phi(\alpha)| \leqq K\alpha^{-l-1}, \quad \alpha \geqq 1,$$

where K is a suitably chosen positive constant.

The first inequality is obvious; the second we demonstrate by contour integration in the complex ϕ-plane. We may describe the contour as the boundary of the simply-connected closed region obtained from the rectangle with vertices $-1, +1, +1-\sigma i, -1-\sigma i, \sigma>0$, by the introduction of a cut along the negative axis of imaginaries from $-i$ to $-\sigma i$. In this region the function $(1-\phi^4)^l e^{-i\alpha\phi}$ is analytic, except at $-1, -i, +1$, and continuous; at the origin it has the value 1. By Cauchy's theorem we have

$$\left(\int_{-1}^{+1} + \int_{+1}^{+1-\sigma i} + \int_{1-\sigma i}^{-\sigma i} + \int_{-\sigma i}^{-i} + \int_{-i}^{-\sigma i} + \int_{-\sigma i}^{-1-\sigma i} + \int_{-1-\sigma i}^{-1}\right)(1-\phi^4)^l e^{-i\alpha\phi}d\phi = 0.$$

On writing $\phi = 1 - it/\alpha$ we find

$$\left|\int_1^{1-\sigma i}(1-\phi^4)^l e^{-i\alpha\phi}d\phi\right| = \left|\frac{1}{\alpha}\int_0^{\sigma\alpha}(1-(1-it/\alpha)^4)^l e^{-i\alpha-t}dt\right|$$

$$\leqq \alpha^{-l-1}\int_0^\infty t^l(4+6t+4t^2+t^3)^l e^{-t}dt \leqq K\alpha^{-l-1}/4,$$

for $\alpha \geqq 1, \sigma > 1$. Similarly, if we set $\phi = t - \sigma i$ we obtain

$$\left|\int_{1-\sigma i}^{-\sigma i}(1-\phi^4)^l e^{-i\alpha\phi}d\phi\right| = \left|\int_0^1(1-(t-\sigma i)^4)^l e^{-\sigma\alpha}e^{-i\alpha t}dt\right|$$

$$\leqq e^{-\sigma}(1+(1+\sigma)^4)^l \to 0$$

as $\sigma \to \infty$, for $\alpha \geqq 1$. Likewise by the substitution $\phi = -i - it/\alpha$ we see that

$$\left|\int_{-\sigma i}^{-i}(1-\phi^4)^l e^{-i\alpha\phi}d\phi\right| = \left|\frac{1}{\alpha}\int_0^{(\sigma-1)\alpha}(1-(1+t/\alpha)^4)^l e^{-\alpha-t}dt\right|$$

$$\leqq \alpha^{-l-1}\int_0^\infty t^l(4+6t+4t^2+t^3)^l e^{-t}dt \leqq K\alpha^{-l-1}/4,$$

for $\alpha \geqq 1, \sigma > 1$. Hence, on allowing σ to become infinite in these three terms and on treating the last three in the same manner, we find

$$|\Phi(\alpha)| \leqq K\alpha^{-l-1}, \quad \alpha \geqq 1,$$

as we were to show.

We can now establish

Theorem XIV. *The Fourier series for an arbitrary summable function is summable* (C, l), $l>0$, *to that function almost everywhere.** *If the function is continuous, then the summability is uniform,* $0<a\leqq x\leqq b<1$.†

Taking the differential system $u'+\lambda u=0$, $u(0)-u(1)=0$ we are to show that if $f(x)$ is summable

$$\lim_{\nu\to\infty}\frac{1}{2\pi i}\int_0^1\int_{C_\nu} f(y)\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l G(x,y\,;\lambda)\,d\lambda\,dy = f(x)$$

almost everywhere on $(0, 1)$. We adopt the notations

$$F(z) = f(z+x) + f(-z+x) - 2f(x),\quad \mathfrak{F}(z) = \int_0^z |F(z)|\,dz.$$

We choose any value of x for which

$$\lim_{z\to 0}\mathfrak{F}(z)/z = 0,\quad 0<x<1.$$

By a theorem of Lebesgue these values of x form a set of measure one.‡ Theorems IV, VII, VIII show us that the only significant contributions of the integral are

$$\begin{aligned}
&\frac{1}{2\pi i}\int_0^1\int_{\gamma_1} f(y)\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l\{e^{\lambda(y-x)}\,;\,0\}\,d\lambda\,dy\\
&+\frac{1}{2\pi i}\int_0^1\int_{\gamma_2} f(y)\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l\{0\,;\,e^{\lambda(y-x)}\}\,d\lambda\,dy\\
&\quad=\frac{1}{2\pi}\int_0^x\int_{-\pi/2}^{+\pi/2} f(y)(1-e^{4i\theta})^l\Lambda_\nu e^{\Lambda_\nu e^{i\theta}(y-x)}e^{i\theta}\,d\theta\,dy\\
&\qquad-\frac{1}{2\pi}\int_x^1\int_{\pi/2}^{3\pi/2} f(y)(1-e^{4i\theta})^l\Lambda_\nu e^{\Lambda_\nu e^{i\theta}(y-x)}e^{i\theta}\,d\theta\,dy,
\end{aligned}$$

where we have recalled the fact that $\lambda=\Lambda_\nu e^{i\theta}$ on γ. Henceforth we shall let Λ be a continuous variable. In the first integral we set $\phi=-ie^{i\theta}$, $z=x-y$, and in the second $\phi=-ie^{i\theta}$, $z=y-x$. There results

$$\frac{1}{2\pi}\int_0^x f(-z+x)\Lambda\Phi(\Lambda z)\,dz+\frac{1}{2\pi}\int_0^{1-x} f(z+x)\Lambda\Phi(\Lambda z)\,dz.$$

---

* Hardy, Proceedings of the London Mathematical Society, ser. 2, vol. 12 (1912), pp. 365–376.

† Chapman, Proceedings of the London Mathematical Society, ser. 2, vol. 9 (1911), pp. 369–409.

‡ Lebesgue, *Leçons sur les Séries Trigonométriques*, Paris, 1905, §50, p. 96.

If we take $A \leqq x, 1-x$ we have as $\Lambda \to \infty$

$$\left| \int_A^x f(-z+x)\Lambda\Phi(\Lambda z)dz \right| \leqq K\Lambda^{-l}A^{-l-1}\int_A^x | f(-z+x) | dz \to 0,$$

$$\left| \int_A^{1-x} f(z+x)\Lambda\Phi(\Lambda z)dz \right| \leqq K\Lambda^{-l}A^{-l-1}\int_A^{1-x} | f(z+x) | dz \to 0.$$

Furthermore, we see that

$$\frac{1}{2\pi}\int_0^A \Lambda\Phi(\Lambda z)dz = \frac{1}{\pi}\int_0^1\int_0^A (1-\phi^4)^l\Lambda \cos \Lambda z\phi dz d\phi = \frac{1}{\pi}\int_0^1 (1-\phi^4)^l \frac{\sin \Lambda A\phi}{\phi} d\phi.$$

The last integral, essentially the well known Dirichlet integral, has the limit $\frac{1}{2}$. To prove our theorem, then, it is sufficient to show that

$$\lim_{\Lambda\to\infty} \frac{1}{2\pi}\int_0^A F(z)\Lambda\Phi(\Lambda z)dz = 0.$$

We can establish this by following Hardy's discussion of a precisely similar integral. In his treatment, the essential properties of the function corresponding to our function Φ are those which we enumerated in Lemma VII. This theorem was demonstrated by Lebesgue in the case $l=1$, and by Hardy in the more general form. It is not difficult to see that when $0<a\leqq x\leqq b<1$ and $f(x)$ is continuous, the steps of the proof can be carried through in the same manner, with the additional result that the convergence is uniform in x because of the fact that

$$\lim_{z\to 0} \mathfrak{F}(z)/z = 0 \quad \text{uniformly,} \quad 0 < a \leqq x \leqq b < 1.$$

This result, due to Fejér for $l=1$, was established in the general case by Chapman.

In our preceding theorems we have compared series only on an interval (a, b) completely interior to the interval $(0, 1)$ on which they are defined. If we wish to consider the entire interval $(0, 1)$ we find it possible to do so when we compare the series arising from differential systems sufficiently alike. It is sufficient to restrict the boundary conditions so that $\alpha_1=\bar{\alpha}_1, \cdots, \alpha_n=\bar{\alpha}_n, \beta_1=\bar{\beta}_1, \cdots, \beta_n=\bar{\beta}_n, k_1=\bar{k}_1, \cdots, k_n=\bar{k}_n$, as we shall see presently. We first introduce new lemmas which serve as useful tools in studying the questions thus raised.

LEMMA VIII. *On any region S'*

$$\frac{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]} = [1] \quad \textit{if} \quad \theta_0 = \bar{\theta}_0, \theta_1 = \bar{\theta}_1.$$

On S_1' we see that

$$([\theta_0]+e^{\rho\omega_\mu}[\theta_1])-([\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]) = [0]+e^{\rho\omega_\mu}[0] = [0],$$

whence

$$\frac{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]} = 1+\frac{[0]}{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]} = [1].$$

Similarly, on S_2', we find

$$\frac{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]} = \frac{[\theta_0]e^{-\rho\omega_\mu}+[\theta_1]}{[\bar{\theta}_0]e^{-\rho\omega_\mu}+[\bar{\theta}_1]} = [1].$$

LEMMA IX. *The integrals*

$$\int_{\gamma_1}\int_\alpha^x e^{\rho\omega_i(1-y)}\frac{m(x,y,\rho)}{\rho}\,dyd\rho \qquad (i=1,\cdots,\mu),$$

$$\int_{\gamma_1}\int_\alpha^x e^{-\rho\omega_i y}\frac{m(x,y,\rho)}{\rho}\,dyd\rho \qquad (i=\mu+1,\cdots,n),$$

$$\int_{\gamma_2}\int_\alpha^x e^{\rho\omega_i(1-y)}\frac{m(x,y,\rho)}{\rho}\,dyd\rho \qquad (i=1,\cdots,\mu-1),$$

$$\int_{\gamma_2}\int_\alpha^x e^{-\rho\omega_i y}\frac{m(x,y,\rho)}{\rho}\,dyd\rho \qquad (i=\mu,\cdots,n),$$

where $|m(x,y,\rho)|\leqq M$ *for all* x *and* y *on* (0, 1) *and for all* ρ *of sufficiently large absolute value on* S_1' *or* S_2', *as the case may be, converge uniformly to zero as* $R\to\infty$, $0\leqq x\leqq 1$, $0\leqq\alpha\leqq 1$.

Leaving aside for the moment the integrals for which $i=\mu$, we may take as typical the following procedure:

$$\left|\int_{\gamma_1}\int_\alpha^x e^{\rho\omega_i(1-y)}\frac{m}{\rho}\,dyd\rho\right| \leqq \frac{\pi M}{2n}\left|\int_\alpha^x e^{-R\beta(1-y)}dy\right|$$

$$= \frac{\pi M}{2\beta nR}\left|e^{-R\beta(1-x)}-e^{-R\beta(1-\alpha)}\right| \leqq \frac{\pi M}{\beta nR}\to 0.$$

In the two integrals where $i=\mu$, we operate as follows:

$$\left|\int_{\gamma_1}\int_\alpha^x e^{\rho\omega_\mu(1-y)}\frac{m}{\rho}\,dy\,d\rho\right| \leqq M\left|\int_\alpha^x\int_0^{\pi/2n} e^{-R\phi(1-y)/2}d\phi dy\right|$$

$$= M\left|\int_\alpha^x \frac{1-e^{-\pi R(1-y)/4n}}{R(1-y)}dy\right| \leqq \frac{M}{R}\left|\int_{(1-\alpha)R}^{(1-x)R}\frac{1-e^{-\pi\xi/4n}}{\xi}d\xi\right|$$

$$\leqq \frac{2M}{R}\int_0^R\frac{1-e^{-\pi\xi/4n}}{\xi}d\xi\to 0.$$

The second integral with $i=\mu$ is discussed in an analogous manner.

We can now prove

THEOREM XV. *If G and $\overline{G}$ are the Green's functions for any two regular differential systems of order $n=2\mu-1$ for which $\alpha_1=\bar{\alpha}_1, \cdots, \alpha_n=\bar{\alpha}_n$, $\beta_1=\bar{\beta}_1, \cdots, \beta_n=\bar{\beta}_n$, $k_1=\bar{k}_1, \cdots, k_n=\bar{k}_n$, then for any region S' the integral*

$$\int_\gamma n\rho^{n-1}(G(x,y\,;\,\rho^n)-\overline{G}(x,y\,;\,\rho^n))d\rho$$

is bounded for all γ and for all x and y on $(0, 1)$.

We must consider separately the integrals over γ_1 and γ_2.

Case I. ρ on S_1'. We refer to the asymptotic forms of $n\rho^{n-1}G$ and $n\rho^{n-1}\overline{G}$ given in Theorem VII. If in the form for $n\rho^{n-1}G$ we expand the determinant Δ_1 we are led to a sum of terms each of which is the product of a certain number of exponential functions, bounded for all ρ on S_1 and for all x and y on $(0, 1)$, multiplied by an expression of the type $[A]/\{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]\}$. In $n\rho^{n-1}\overline{G}$ the term containing the corresponding exponentials has the coefficient $[\overline{A}]/\{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]\}$ where by hypothesis $A=\overline{A}$, $\theta_0=\bar{\theta}_0$, $\theta_1=\bar{\theta}_1$. If we subtract the corresponding terms, the resulting coefficient becomes, by Lemma VIII,

$$\frac{[A]}{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}-\frac{[\overline{A}]}{[\bar{\theta}_0]+e^{\rho\omega_\mu}[\bar{\theta}_1]}=\frac{[A]-[\overline{A}]}{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}=\frac{[0]}{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]}=[0]$$

for all ρ on S_1' and all x and y on $(0, 1)$. Now each product of exponentials from the determinants Δ_1, $\overline{\Delta}_1$ contains exactly one of the functions

$$e^{\rho\omega_1(1-y)},\ \cdots,\ e^{\rho\omega_\mu(1-y)},\ e^{-\rho\omega_{\mu+1}y},\ \cdots,\ e^{-\rho\omega_n y}.$$

If we group the terms in the difference $n\rho^{n-1}(G-\overline{G})$ according to the exponentials in y which they contain, we see that we can write

$$\begin{aligned}n\rho^{n-1}(G-\overline{G})&=\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho};+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\frac{m_i(x,y,\rho)}{\rho}\right\}\\&\quad+\sum_{i=1}^{i=\mu}e^{\rho\omega_i(1-y)}\frac{m_{n+i}(x,y,\rho)}{\rho}+\sum_{i=\mu+1}^{i=n}e^{-\rho\omega_i y}\frac{m_{n+i}(x,y,\rho)}{\rho}\\&=\frac{m(x,y,\rho)}{\rho},\end{aligned}$$

where the functions m are bounded for all ρ on S_1' and all x and y on $(0, 1)$. The application of Lemma III to the last expression proves that

$$\int_{\gamma_1}n\rho^{n-1}(G(x,y;\rho^n)-\overline{G}(x,y\,;\rho^n))d\rho$$

is bounded for all γ_1 on S_1' and for all x and y on $(0, 1)$.

Case II. ρ on S_2'. We now find for the difference $n\rho^{n-1}(G-\overline{G})$ an expression like that of Case I, different only in having the sums extended from $i=1$ to $i=\mu-1$ and from $i=\mu$ and to $i=n$ respectively. The desired result for the integral over γ_2 is then established at once; and the theorem follows at once.

THEOREM XVI. *Under the hypotheses of Theorem* XV,

$$\lim_{R\to\infty}\int_\alpha^\beta\int_\gamma n\rho^{n-1}(G(x,y\,;\,\rho^n)-\overline{G}(x,y\,;\,\rho^n))d\rho dy=0$$

uniformly for all x on (0, 1).

We take the integral over (α, β) as the sum of integrals over (α, x) and (x, β). Of these we consider but one.

Case I. ρ on S_1'. From the expressions obtained in Theorem XV we see that

$$\int_\alpha^x\int_{\gamma_1} n\rho^{n-1}(G-\overline{G})d\rho dy$$

$$=\int_\alpha^x\int_{\gamma_1}\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho};+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho}\right\}d\rho dy$$

$$+\int_\alpha^x\int_{\gamma_1}\left(\sum_{i=1}^{i=\mu}e^{\rho\omega_i(1-y)}\frac{m_{n+i}}{\rho}+\sum_{i=\mu+1}^{i=n}e^{-\rho\omega_i y}\frac{m_{n+i}}{\rho}\right)d\rho dy\to 0,$$

uniformly, $0\leqq x\leqq 1$, $0\leqq\alpha\leqq 1$, as $R\to\infty$, by Lemmas VI and IX.

Case II. ρ on S_2'. We use the work of Theorem XV and Lemmas VI and IX exactly as in the preceding case. The double integral over γ and (α, β) therefore behaves as described.

THEOREM XVII. *Under the hypotheses of Theorem* XV,

$$\lim_{R\to\infty}\int_0^1 f(y)\int_\gamma n\rho^{n-1}(G(x,y\,;\,\rho^n)-\overline{G}(x,y\,;\,\rho^n))d\rho dy=0,$$

$$\lim_{\Lambda\to\infty}\int_0^1 f(y)\int_C (G(x,y\,;\,\lambda)-\overline{G}(x,y\,;\,\lambda))d\lambda dy=0$$

uniformly for all x on (0, 1), ***for any summable function*** $f(x)$.

We combine Theorems IV, XV, XVI to obtain the limit of the first integral. If we take arcs γ on two adjacent sectors S' and transfer the results obtained for the integrals over them to the λ-plane, we find the second limit.

As Theorems IX and XII resulted in the first part of Theorem XIII, so Theorems XII and XVII lead to

THEOREM XVIII. *The term-by-term difference of two Birkhoff series of order $n=2\mu-1$ for any summable function $f(x)$ converges uniformly to zero, $0\leqq x\leqq 1$, provided that the boundary conditions of the two differential systems defining the series can be reduced to normal forms for which $\alpha_1=\bar{\alpha}_1, \cdots, \alpha_n=\bar{\alpha}_n, \beta_1=\bar{\beta}_1, \cdots, \beta_n=\bar{\beta}_n, k_1=\bar{k}_1, \cdots, k_n=\bar{k}_n$. Such Birkhoff series are equivalent on* (0, 1). *The differential system*

$$\begin{aligned}
&u^{(n)} + \lambda u = 0, && n = 2\mu - 1,\\
&\alpha_1 u^{(k_1)}(0) + \beta_1 u^{(k_1)}(1) = 0, && n - 1 \geqq k_1 \geqq k_2 \geqq \cdots \geqq k_n, k_{i+2} > k_i,\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot && \\
&\alpha_n u^{(k_n)}(0) + \beta_n u^{(k_n)}(1) = 0, &&
\end{aligned}$$

where α, β, k satisfy the conditions for regularity, can therefore be regarded as typical of regular differential systems of order $n=2\mu-1$ in the consideration of expansion problems with respect to the class of all summable functions.

This theorem is the generalization to the case $n=2\mu-1$ of the leading results of Haar and Walsh mentioned in the introductory paragraph. We shall apply it to the discussion of Birkhoff series at $x=0$ and $x=1$.

We shall introduce two determinants $\mathcal{D}_1(x, y, \rho)$ and $\mathcal{D}_2(x, y, \rho)$ which can be obtained from the two determinants Δ_1 and Δ_2 described under Theorem VII by replacing each term of the form $[A]$ by the dominant constant A, and deleting the terms in $e^{\rho\omega_\mu}$. We can then prove an interesting intermediate theorem.

THEOREM XIX. *If $f(x)$ is any summable function on* (0, 1) *and if G is the Green's function for a regular differential system of order $n=2\mu-1$, then*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_1}\Bigl(n\rho^{n-1}G(x,y\,;\,\rho^n) - \Bigl\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}\omega_i\,;+\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}\omega_i\Bigr\}$$
$$-\frac{1}{\theta_0}\mathcal{D}_1(x,y,\rho)\Bigr)d\rho dy = 0,$$
$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_2}\Bigl(n\rho^{n-1}G(x,y\,;\,\rho^n) - \Bigl\{-\sum_{i=1}^{i=\mu-1} e^{\rho\omega_i(x-y)}\omega_i\,;+\sum_{i=\mu}^{i=n} e^{\rho\omega_i(x-y)}\omega_i\Bigr\}$$
$$-\frac{1}{\theta_1}\mathcal{D}_2(x,y,\rho)\Bigr)d\rho dy = 0,$$

uniformly, $0\leqq x\leqq 1$.

By Theorem XVIII, it is sufficient to consider the typical differential system there described.

First integral. On substituting $u_i = e^{\rho\omega_i x}$, $v_i = -\omega_i e^{-\rho\omega_i y}/n\rho^{n-1}$ in the explicit formula for the Green's function, we obtain for ρ on S_1' the asymptotic form given in Theorem VII, where now any expression $[A]$ can be written as the sum of the constant A and a number of exponential terms formed by taking products of $e^{\rho\omega_1}, \cdots, e^{\rho\omega_{\mu-1}}, e^{\rho\omega_{\mu+1}}, \cdots, e^{\rho\omega_n}$. The expressions $[A]$ in the bracket terms and in the first row of the determinant Δ_1 reduce to the constant A. Since we can put

$$e^{\rho\omega_i} = e^{\rho\omega_\mu}e^{\rho\omega_i-\rho\omega_\mu}, \quad |e^{\rho\omega_i-\rho\omega_\mu}| = e^{R(\rho\omega_i)-R(\rho\omega_\mu)} \leqq 1 \quad (i = 1, \cdots, \mu-1),$$
$$e^{-\rho\omega_i} = e^{\rho\omega_\mu}e^{-\rho\omega_i-\rho\omega_\mu}, \quad |e^{-\rho\omega_i-\rho\omega_\mu}| = e^{-R(\rho\omega_i)-R(\rho\omega_\mu)} \leqq 1 \quad (i = \mu+1, \cdots, n),$$

for all ρ on S_1', we find $[A] = A + e^{\rho\omega_\mu}m(\rho)$. If now we expand the determinant Δ_1 we see that

$$\Delta_1 = \mathcal{D}_1(x,y,\rho) + e^{\rho\omega_\mu}\sum_{i=1}^{i=\mu} e^{\rho\omega_i(1-y)}m_i(x,\rho) + e^{\rho\omega_\mu}\sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}m_i(x,\rho).$$

Slight manipulation shows that

$$\frac{1}{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]} = \frac{1}{\theta_0} - \frac{e^{\rho\omega_\mu}m(\rho)}{[\theta_0]+e^{\rho\omega_\mu}[\theta_1]} = \frac{1}{\theta_0} - e^{\rho\omega_\mu}M(\rho)$$

for all ρ on S_1'. Therefore we find

$$n\rho^{n-1}G - \left\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}\omega_i\,;+\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}\omega_i\right\} - \frac{1}{\theta_0}\mathcal{D}_1$$
$$= e^{\rho\omega_\mu}\left(\sum_{i=1}^{i=\mu} e^{\rho\omega_i(1-y)}M_i(x,\rho) + \sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}M_i(x,\rho)\right) = e^{\rho\omega_\mu}M(x,y,\rho)$$

where the functions M are uniformly bounded for all x and y on $(0, 1)$ and for all ρ on S_1'. The first integral can now be brought under Theorem IV. For

$$\left|\int_{\gamma_1} e^{\rho\omega_\mu}M(x,y,\rho)d\rho\right| \leqq M\int_0^{\pi/2n} e^{-R\phi/2}Rd\phi = 2M(1-e^{-\pi R/4n}) \leqq 2M,$$

while

$$\left|\int_\alpha^\beta\int_{\gamma_1} e^{\rho\omega_\mu}\left(\sum_{i=1}^{i=\mu} e^{\rho\omega_i(1-y)}M_i(x,\rho) + \sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}M_i(x,\rho)\right)d\rho dy\right|$$
$$= \left|\int_{\gamma_1} e^{\rho\omega_\mu}\left(\sum_{i=1}^{i=\mu} M_i\frac{e^{\rho\omega_i(1-\alpha)}-e^{\rho\omega_i(1-\beta)}}{\omega_i} + \sum_{i=\mu+1}^{i=n} M_i\frac{e^{-\rho\omega_i\alpha}-e^{-\rho\omega_i\beta}}{\omega_i}\right)\frac{d\rho}{\rho}\right|$$
$$\leqq M'\int_0^{\pi/2n} e^{-R\phi/2}d\phi \to 0, \qquad 0 \leqq x \leqq 1.$$

The theorem is therefore established for the first integral.

Second integral. By proceeding as in the case of the first integral, we show that the integrand in the second has entirely similar properties, due account being taken of the behavior of $\rho\omega_\mu$ on S_2'. The theorem is then easily proved for this integral.

In these facts, then, we see the expansion problem set by Birkhoff stripped of all asymptotic encumbrances and expressed explicitly as a problem involving only the elementary functions in linear combination.

On setting $x=0$, $x=1$, in the integrals of the preceding theorem we can effect still further reductions. We denote by D_1^0, D_1^1, D_2^0, D_2^1 determinants obtained from $\mathcal{D}_1(0, y, \rho)$, $\mathcal{D}_1(1, y, \rho)$, $\mathcal{D}_2(0, y, \rho)$, $\mathcal{D}_2(1, y, \rho)$ respectively by appropriate modifications of the first rows; more exactly, in the last four determinants those terms of the first rows explicitly involving the variable ρ as an exponential are replaced by zeros.

THEOREM XX. *The limits described in Theorem* XIX *are still valid when x is set equal to* 0 *or* 1 *and the determinants $\mathcal{D}$ are replaced by the corresponding determinants D. For example,*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_1}\Big(n\rho^{n-1}G(0,y\,;\rho^n) - \sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}\omega_i - \frac{1}{\theta_0}D_1^0(y,\rho)\Big)d\rho dy = 0.$$

By Theorem XIX it is sufficient to show that

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_1}(\mathcal{D}_i(k,y,\rho) - D_i^k(y,\rho))d\rho dy = 0$$

for $i=1$, 2 and $k=0$, 1. Since we have, by reasoning like that used in the proof of Theorem XIX,

$$\mathcal{D}_1(k,y,\rho) - D_1^k(y,\rho) = e^{\rho\omega_\mu}m_k(y,\rho)$$

where the functions m are uniformly bounded, $0\leqq y\leqq 1$, for ρ on S_1, $k=0$, 1; and since we can obtain similar expressions for the difference $\mathcal{D}_2(k, y, \rho) - D_2^k(y, \rho)$ for $k=0$, 1, we are able to complete the proof exactly as in Theorem XIX.

To apply Theorem XX to the study of the convergence of Birkhoff series at $x=0$ and $x=1$ we demonstrate the three lemmas which follow. These lemmas are generalisations of standard theorems concerning the Dirichlet integral.

LEMMA X. *If $f(x)$ is summable on* (0, 1) *and if $\lambda_1=e^{i\theta_1}$, and $\lambda=e_2^{i\theta_2}$, where θ_1 and θ_2 are two distinct constants on the range $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, then*

$$\lim_{R\to\infty}\int_\epsilon^1 f(y)\,\frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy = 0, \quad \epsilon > 0.$$

Since the first factor of the integrand, $f(y)/y$, is summable on $(\epsilon, 1)$ and since the second, $e^{-R\lambda_1 y}-e^{-R\lambda_2 y}$, satisfies the sufficient conditions of Theorem IV, as may easily be shown, the lemma is apparent.

LEMMA XI. *If $\phi(x)$ is summable on* $(0, 1)$ *and of bounded variation on* $(0, \epsilon)$, *then*

$$\lim_{R\to\infty}\int_0^1 \phi(y)\,\frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy = \phi(+0)(\theta_2 - \theta_1)i.$$

We show that the function $(e^{-R\lambda_1 y}-e^{-R\lambda_2 y})/y$ has two properties which are the essential elements in well known proofs of this lemma in the case of the Dirichlet integral.

In the first place, we show that

$$\lim_{R\to\infty}\int_0^1 \frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy = (\theta_2 - \theta_1)i.$$

In the plane of the complex variable t we determine a contour consisting of the rectilinear segments $(r\lambda_1, R\lambda_1)$ and $(r\lambda_2, R\lambda_2)$ and two circular arcs c and C of radii r and R respectively described about the origin and lying on the right half-plane. Then

$$\lim_{R\to\infty}\int_0^1 (e^{-R\lambda_1 y} - e^{-R\lambda_2 y})\,\frac{dy}{y} = \lim_{R\to\infty}\ \lim_{r\to 0}\left(\int_{r\lambda_1}^{R\lambda_1}\frac{e^{-t}}{t}\,dt - \int_{r\lambda_1}^{R\lambda_2}\frac{e^{-t}}{t}\,dt\right).$$

By using the contour described and applying Cauchy's theorem, we can replace the last limit by

$$\lim_{r\to\infty}\int_c \frac{e^{-t}}{t}\,dt - \lim_{R\to\infty}\int_C \frac{e^{-t}}{t}\,dt = (\theta_2 - \theta_1)i - \lim_{R\to\infty}\int_C \frac{e^{-t}}{t}\,dt = (\theta_2 - \theta_1)i.$$

The second property needed is that the integral

$$\int_a^b \frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy$$

is a uniformly bounded function of R for all a and b on $(0, 1)$, no matter how a and b may vary with R. Clearly it is sufficient to show this for the special case $a=0$. On writing $B=Rb$, we find

$$\int_0^b \frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy = (\theta_2 - \theta_1)i - \int_{B\lambda_1}^{B\lambda_2}\frac{e^{-t}dt}{t},$$

whence the result is apparent.

It is now a simple matter to demonstrate the lemma.*

LEMMA XII. *If $\phi(x)$ is summable on* $(0, 1)$ *and if*

$$\Phi(x) = \frac{1}{x}\int_0^x \phi(x)dx$$

is of bounded variation on $(0, \epsilon)$, *then*

$$\lim_{R\to\infty}\int_0^1 \phi(y)\,\frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy = \Phi(+0)(\theta_2 - \theta_1)i.$$

The proof can be carried out along the lines indicated by de la Vallée Poussin in a paper in which he established this result for the Dirichlet integral.† As he points out, this lemma includes Lemma XII as a special case.

We next prove

THEOREM XXI. *If $f(x)$ is summable on* $(0, 1)$, *then the convergence at $x=0$ or at $x=1$ of the Birkhoff series of order $n=2\mu-1$ for this function is independent of the nature of $f(x)$ outside arbitrarily small neighborhoods of the points $x=0$ and $x=1$.*

By Theorem XX the convergence of the series in question depends upon the behavior of certain integrals which we can write out explicitly and reduce to the type discussed in Lemmas X–XII. These integrals arise from the last columns of the determinants D. We use $\lambda_1, \lambda_2, \theta_1, \theta_2$ to denote any constants satisfying the relations laid down in Lemma X. Then, by fundamental properties of $\rho\omega_1, \cdots, \rho\omega_n$ on S_1, S_2, we find that the integrals in question can be put in the forms

$$\int_0^1 f(y)\int_{\gamma_1}\omega_k e^{\rho\omega_k(1-y)}d\rho dy = \int_0^1 f(1-y)\int_{\gamma_1}\omega_k e^{\rho\omega_k y}d\rho dy$$
$$= \int_0^1 f(1-y)\frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}dy$$

where $\theta_2 - \theta_1 = -\pi/2n,\ k = 1, \cdots, \mu$;

$$\int_0^1 f(y)\int_{\gamma_1}\omega_k e^{-\rho\omega_k y}d\rho dy = \int_0^1 f(y)\,\frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\,dy$$

---

* Hobson, *The Theory of Functions of a Real Variable*, Cambridge, 1907, ed. 1, §450.

† de la Vallée Poussin, Rendiconti del Circolo Matematico di Palermo, vol. 31 (1911), pp. 296–299.

where $\theta_2 - \theta_1 = +\pi/2n,\ k = \mu+1, \cdots, n$;

$$\int_0^1 f(y)\int_{\gamma_2} \omega_k e^{\rho\omega_k(1-y)}d\rho dy = \int_0^1 f(1-y)\,\frac{e^{-R\lambda_1 y}-e^{-R\lambda_2 y}}{y}\,dy$$

where $\theta_2 - \theta_1 = -\pi/2n,\ k = 1, \cdots, \mu-1$;

$$\int_0^1 f(y)\int_{\gamma_2} \omega_k e^{-\rho\omega_k y}d\rho dy = \int_0^1 f(y)\,\frac{e^{-R\lambda_1 y}-e^{-R\lambda_2 y}}{y}\,dy$$

where $\theta_2-\theta_1=+\pi/2n$, $k=\mu, \cdots, n$. It is immediately evident from Lemma X that the behavior of these integrals, and therefore of the series for $f(x)$ at $x=0$ or at $x=1$, is independent of the nature of $f(x)$ except near these points.

The use of the integrals computed in the preceding paragraph and of Lemma XII enables us to state

THEOREM XXII. *If $\phi(x)$ is summable on $(0, 1)$, and if*

$$\Phi_1(x) = \frac{1}{x}\int_0^x \phi(x)dx, \quad \Phi_2(x) = \frac{1}{x}\int_0^x \phi(1-x)dx$$

are of bounded variation near $x=0$, then

$$\lim_{R\to\infty}\frac{1}{2\pi\sqrt{-1}}\int_0^1 \phi(y)\int_{\gamma_i} n\rho^{n-1}G(k,y\,;\,\rho^n)d\rho dy = A_{ik}\Phi_1(+0) + B_{ik}\Phi_2(+0)$$

where A_{ik} and B_{ik} are constants depending only upon the boundary conditions, for $i=1, 2$, $k=0, 1$. For example,

$$A_{10}\Phi_1(+0) + B_{10}\Phi_2(+0) \equiv \frac{\mu-1}{4n}\Phi_1(+0)$$
$$+\frac{1}{4n\theta_0}\begin{vmatrix} 1 & \cdots 1 & 0 & \cdots 0 & 0 \\ \alpha_1\omega_1^{k_1} & \cdots \alpha_1\omega_\mu^{k_1} & \beta_1\omega_{\mu+1}^{k_1} & \cdots \beta_1\omega_n^{k_1} & -\beta_1 a_2\sum_{i=1}^{i=\mu}\omega_i^{k_1} - \alpha_1 a_1 \sum_{i=\mu+1}^{i=n}\omega_i^{k_1} \\ \cdot\ \cdot & \cdots\ \cdot\ \cdot & \cdot\ \cdot & \cdots\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \alpha_n\omega_1^{k_n} & \cdots \alpha_n\omega_\mu^{k_n} & \beta_n\omega_{\mu+1}^{k_n} & \cdots \beta_n\omega_n^{k_n} & -\beta_n a_2\sum_{i=1}^{i=\mu}\omega_i^{k_n} - \alpha_n a_1 \sum_{i=\mu+1}^{i=n}\omega_i^{k_n} \end{vmatrix},$$

where $a_1=\Phi_1(+0)$, $a_2=\Phi_2(+0)$. The other constants may be explicitly determined in the same manner by reference to Theorem XX *and the forms of the determinants D.*

This theorem is a generalization of the result announced by Birkhoff for the case $n=2\mu$*.

We shall make no further applications of Theorem XX. In Theorem IV the sufficient conditions given are also necessary; by combining them with Theorem XX we find necessary conditions that

$$\lim_{\Lambda\to\infty}\int_0^1\int_C f(y)\big(G(x,y\,;\lambda)-\bar{G}(x,y\,;\lambda)\big)d\lambda dy=0$$

uniformly, $0\leqq x\leqq 1$. It seems as though these conditions in their simplest form should involve only the constants α, $\bar{\alpha}$, β, $\bar{\beta}$, k, $\bar{k}$, but we have been unable to ascertain the accuracy of this conjecture.

V. The adjoint series for $n=2\mu-1$

We shall next take up the expansion problem which comes to us from the integral

$$\frac{1}{2\pi i}\int_0^1 f(y)\int_{C_\nu} G(y,x\,;\lambda)d\lambda dy,$$

where $G(x,y;\lambda)$ is the Green's function associated with a regular differential system of order $n=2\mu-1$. If this differential system admits an adjoint, in which the parameter appears as $-\lambda$ instead of $+\lambda$, the Green's function for this adjoint system is precisely $G(y,x;\lambda)$. Since the adjoint system is essentially a regular differential system of order $n=2\mu-1$ of the type discussed in §IV, the work of that section can be applied directly. Since the differential systems which we are discussing are too general to admit adjoints except in restricted cases, we are faced with a new expansion problem. The series obtained we call the adjoint series of order $n=2\mu-1$, for obvious reasons. Fortunately it is possible to throw the problem into a form which presents such similarities to that discussed in §IV that we do not need to do more than paraphrase the principal theorems demonstrated hitherto.

In fact, if we make the substitutions $\bar{x}=1-x$, $\bar{y}=1-y$, $\bar{f}(\bar{y})=f(1-\bar{y})$, we find at once

$$\int_0^1 f(y)\int_{C_\nu} G(y,x\,;\lambda)d\lambda dy=\int_0^1\bar{f}(\bar{y})\int_{C_\nu} G(1-\bar{y},1-\bar{x}\,;\lambda)d\lambda d\bar{y}.$$

Then, by interchanging x and y in the asymptotic forms of Theorem VII and subsequently making the substitutions described, we obtain asymptotic

* Birkhoff, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), pp. 125–126.

forms for $n\rho^{n-1}G(1-\bar{y},\ 1-\bar{x};\ \rho^n)$ which are to all intents and purposes the same as those studied in the preceding section. The fact that in the determinants involved the terms in x and y occur in the last column and the first row respectively instead of vice versa is entirely unessential. Thus for the difference $G(1-\bar{y}, 1-\bar{x};\lambda)-\bar{G}(1-\bar{y}, 1-\bar{x};\lambda)$ where $G(x, y;\lambda)$ and $\bar{G}(x, y;\lambda)$ are the Green's functions for two regular differential systems of order $n=2\mu-1$ we can state theorems analogous to Theorems IX and XVII. For each Green's function there is a theorem parallel to Theorem XII. Then we find the analogues of Theorems XIII and XVIII. Theorems XIX, XX, XXI and XXII likewise have images in the present case, not only by reasoning parallel to that of §IV, but also by the fact that the Green's function can at this stage be regarded as arising from the typical differential system of Theorem XVIII, which admits an adjoint falling under §IV. When these theorems have been phrased in terms of the variables $\bar{x}$ and $\bar{y}$, it is easy to return to the variables x and y.

Thus we have

THEOREM XXIII. *If $f(x)$ is summable on* (0, 1)

$$\lim_{\Lambda\to\infty}\int_0^1 f(y)\int_C (G(y,x\ ;\ \lambda) - \bar{G}(y,x\ ;\ \lambda))d\lambda dy = 0$$

uniformly, $0<a\leqq x\leqq b<1$.

THEOREM XXIV. *If $f(x)$ is summable on* (0, 1) *and if $R_\nu(y, x)$ is the residue of $G(y, x;\lambda)$ at $\lambda=\lambda_\nu$, then*

$$\lim_{\nu\to\infty}\int_0^1 f(y)R_\nu(y,x)dy = 0$$

uniformly, $0\leqq x\leqq 1$.

THEOREM XXV. *On any closed interval* (a, b) *completely interior to* (0, 1) *the term-by-term difference of the adjoint series formed for any summable function $f(x)$ in connection with two regular differential systems of order $n=2\mu-1$ converges uniformly to zero. We say that these adjoint series are equivalent on (a, b); in particular, they are equivalent to Fourier series on (a, b)*

The last remark results from the fact that the Fourier series obtained as special Birkhoff series in § III are not altered by an interchange of x and y.

THEOREM XXVI. *The Green's functions associated with two regular differential systems of order $n=2\mu-1$ for which $\alpha_1=\bar{\alpha}_1, \cdots, \alpha_n=\bar{\alpha}_n$, $\beta_1=\bar{\beta}_1, \cdots, \beta_n=\bar{\beta}_n$, $k_1=\bar{k}_1, \cdots, k_n=\bar{k}_n$, are such that for any summable function $f(x)$*

$$\lim_{\Lambda\to\infty}\int_0^1 f(y)\int_C \left(G(y,x\,;\,\lambda)-\bar{G}(y,x\,;\,\lambda)\right)d\lambda dy = 0$$

uniformly, $0\leqq x\leqq 1$.

THEOREM XXVII. *On the interval* (0, 1) *the term-by-term difference of the adjoint series formed for any summable function $f(x)$ in connection with two regular differential systems of the form described in Theorem* XXVI *converges uniformly to zero. Such adjoint series are equivalent on* (0, 1). *The adjoint series arising from the regular differential system*

$$\begin{aligned} &u^{(n)}+\lambda u=0, && n=2\mu-1,\\ &\alpha_1 u^{(k_1)}(0)+\beta_1 u^{(k_1)}(1)=0, && n-1\geqq k_1\geqq k_2\geqq\cdots\geqq k_n,\ k_{i+2}>k_i,\\ &\cdots\cdots\cdots\\ &\alpha_n u^{(k_n)}(0)+\beta_n u^{(k_n)}(1)=0 \end{aligned}$$

may be regarded as typical, so far as the class of all summable functions is concerned.

THEOREM XXVIII. *If $f(x)$ is summable on* (0, 1) *then*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_1}\Bigg(n\rho^{n-1}G(y,x\,;\,\rho^n)-\Bigg\{\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(y-x)}\omega_i\,;\,-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(y-x)}\omega_i\Bigg\}$$
$$-\frac{1}{\theta_0}\,\mathcal{D}_1(y,x,\rho)\Bigg)d\rho dy=0,$$

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_2}\Bigg(n\rho^{n-1}G(y,x\,;\,\rho^n)-\Bigg\{\sum_{i=\mu}^{i=n} e^{\rho\omega_i(y-x)}\omega_i\,;\,-\sum_{i=1}^{i=\mu-1} e^{\rho\omega_i(y-x)}\omega_i\Bigg\}$$
$$-\frac{1}{\theta_1}\,\mathcal{D}_2(y,x,\rho)\Bigg)d\rho dy=0,$$

uniformly, $0\leqq x\leqq 1$.

We now obtain from the four determinants $\mathcal{D}_1(y,0,\rho)$, $\mathcal{D}_1(y,1,\rho)$, $\mathcal{D}_2(y,0,\rho)$, $\mathcal{D}_2(y,1,\rho)$ respectively four new determinants $\mathfrak{D}_1^0$, $\mathfrak{D}_1^1$, $\mathfrak{D}_2^0$, $\mathfrak{D}_2^1$ analogous to the determinants D introduced in Theorem XX. In any determinant $\mathcal{D}$ the terms in the last column explicitly involving the variable ρ are replaced by zero to give the corresponding determinant $\mathfrak{D}$.

Theorem XXIX. *If $f(x)$ is summable on* (0, 1) *the limits described in Theorem* XXVIII *are valid when x is set equal to* 0 *or* 1 *and the determinants* D *are replaced by the corresponding determinants* $\mathfrak{D}$. *For example,*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_{\gamma_1}\left(n\rho^{n-1}G(y,0\,;\,\rho^n)+\sum_{i=1}^{i=\mu}e^{\rho\omega_i y}\omega_i-\frac{1}{\theta_0}\mathfrak{D}_1{}^0(y,\rho)\right)d\rho dy=0.$$

Theorem XXX. *If $f(x)$ is summable on* (0, 1), *then the convergence at $x=0$ or at $x=1$ of the formal adjoint series of order $n=2\mu-1$ for this function is independent of the nature of $f(x)$ outside arbitrarily small neighborhoods of the points $x=0$ and $x=1$.*

Theorem XXXI. *If $\phi(x)$ is summable on* (0, 1) *and if*

$$\Phi_1(x)=\frac{1}{x}\int_0^x\phi(x)dx,\qquad \Phi_2(x)=\frac{1}{x}\int_0^x\phi(1-x)dx$$

are of bounded variation near $x=0$, then

$$\lim_{R\to\infty}\frac{1}{2\pi\sqrt{-1}}\int_0^1\phi(y)\int_{\gamma_i}n\rho^{n-1}G(y,k\,;\,\rho^n)d\rho dy=\mathfrak{A}_{ik}\Phi_1(+0)+\mathfrak{B}_{ik}\Phi_2(+0)$$

where the constants $\mathfrak{A}$ and $\mathfrak{B}$ depend only on the boundary conditions, $i=1, 2$, $k=0, 1$. In particular,

$$\mathfrak{A}_{10}\Phi_1(+0)+\mathfrak{B}_{10}\Phi_2(+0)\equiv\frac{\mu}{4n}\Phi_1(+0)$$

$$+\frac{1}{4n\theta_1}\begin{vmatrix} a_1 & \cdots & a_\mu & a_{\mu+1} & \cdots & a_n & 0\\ \alpha_1\omega_1{}^{k_1} & \cdots & \alpha_1\omega_\mu{}^{k_1} & \beta_1\omega_{\mu+1}^{k_1} & \cdots & \beta_1\omega_n{}^{k_1} & -\alpha_1\sum_{i=\mu+1}^{i=n}\omega_i{}^{k_1+1}\\ \cdot\ \cdot & \cdots & \cdot\ \cdot & \cdot\ \cdot & \cdots & \cdot\ \cdot & \cdot\ \cdot\ \cdot\\ \alpha_n\omega_1{}^{k_n} & \cdots & \alpha_n\omega_\mu{}^{k_n} & \beta_n\omega_{\mu+1}^{k_n} & \cdots & \beta_n\omega_n{}^{k_n} & -\alpha_n\sum_{i=\mu+1}^{i=n}\omega_i{}^{k_n+1} \end{vmatrix},$$

where $a_i=-\Phi_1(+0)/\omega_i$, $i=1,2,\cdots,\mu$, and $a_i=\Phi_2(+0)/\omega_i$, $i=\mu+1,\cdots,n$. The other constants may be expressed similarly by reference to Theorem XXIX *and the explicit forms of the determinants* $\mathfrak{D}$.

This completes our study of the adjoint series of order $n=2\mu-1$.

VI. The derived series of Birkhoff, $n=2\mu-1$

It is now convenient to restrict our attention to the regular differential system

$$u^{(n)} + \bigstar + p_2 u^{(n-2)} + \cdots + (p_n + \lambda)u = 0,$$
$$W_1(u) = 0, \cdots, W_n(u) = 0,$$

where $p_2, \cdots, p_n$ are real or complex functions of the real variable x, continuous together with their derivatives of all orders on $(0, 1)$. The derived series of Birkhoff are then expressed as the integral

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu} G(x,y\,;\lambda)d\lambda dy \qquad (k = 1,2,\cdots),$$

while a method of summing these series is given by

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\lambda)d\lambda dy, \quad l \geqq 0.$$

If we wish to consider the series only for $k=1, 2, \cdots, K$, the functions $p_2, \cdots, p_n$ do not have to be so heavily restricted; but the complications in statement thus introduced do not yield a proportionate increase in interest.

Lemma XIII. *Under the present hypotheses*

$$\begin{aligned}\frac{\partial^k}{\partial x^k}&\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\lambda)d\lambda dy\\ &= \int_0^x f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}\frac{\partial^k}{\partial x^k}G(x,y\,;\lambda)d\lambda dy\\ &\quad + \int_x^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}\frac{\partial^k}{\partial x^k}G(x,y\,;\lambda)d\lambda dy.\end{aligned}$$

Since

$$\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\lambda)d\lambda$$

is continuous together with its partial derivatives of all orders with respect to x for all x and y on $(0, 1)$ we have

$$\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} Gd\lambda dy = \int_0^1 f(y)\frac{\partial^k}{\partial x^k}\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} Gd\lambda dy.$$

Now on the ranges $0\leqq y<x$, $0\leqq x\leqq 1$, and $x<y\leqq 1$, $0\leqq x\leqq 1$, $G(x,y;\lambda)$ is continuous together with its partial derivatives of all orders with respect to x. Hence, on each of these ranges

$$\frac{\partial^k}{\partial x^k}\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} Gd\lambda dy = \int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}\frac{\partial^k}{\partial x^k}Gd\lambda dy.$$

The lemma follows at once.

Lemma XIV. *The functions $v_1(y), \cdots, v_n(y)$ defined by the identity*

$$\tau_1 v_1(y) + \cdots + \tau_n v_n(y) \equiv \begin{vmatrix} u_1(y) & \cdots u_n(y) \\ \cdot & \cdots \cdot \\ u_1^{(n-2)}(y) & \cdots u_n^{(n-2)}(y) \\ \tau_1 & \cdots \tau_n \end{vmatrix} \div \begin{vmatrix} u_1(y) & \cdots u_n(y) \\ \cdot & \cdots \cdot \\ \cdot & \cdots \cdot \\ u_1^{(n-1)}(y) & \cdots u_n^{(n-1)}(y) \end{vmatrix}$$

have for a given sector S the asymptotic forms

$$v_i(y) \equiv \frac{-\omega_i}{n\rho^{n-1}} e^{-\rho\omega_i y}\left(1 + \sum_{l=1}^{l=m} \frac{B_l(y)}{(\rho\omega_i)^l} + \frac{E_i(y,\rho)}{\rho^{m+1}}\right),$$

$i=1, \cdots, n$, where the functions $B_l(y)$, $l=1, \cdots, m$, are continuous together with their derivatives of all orders on (0, 1), *and the functions E are uniformly bounded for all y on* (0, 1) *and for all ρ of sufficiently great absolute value on S. The functions $B_l(y)$ are independent of the sector S.**

We refer to Theorem III′, from which we obtain

$$u_i^{(k)}(y) = (\rho\omega_i)^k e^{\rho\omega_i y} P_k(\rho\omega_i) + (\rho\omega_i)^k e^{\rho\omega_i y} \frac{E_{ik}(y,\rho)}{\rho^{m+1}}$$

where $P_k(z)$ is a polynomial in $1/z$, with coefficients which are functions of y and with the constant term 1. The denominator determinant in the defining expression for the functions $v_1(y), \cdots, v_n(y)$ is known to be independent of y. Hence we have

$$v_i(y) = \frac{\begin{vmatrix} u_1(y) & \cdots u_{i-1}(y) & u_i(y) & u_{i+1}(y) & \cdots u_n(y) \\ \cdot\ \cdot & \cdots\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot & \cdots\ \cdot\ \cdot \\ u_1^{(n-2)}(y) & \cdots u_{i-1}^{(n-2)}(y) & u_i^{(n-2)}(y) & u_{i+1}^{(n-2)}(y) & \cdots u_n^{(n-2)}(y) \\ 0 & 0 & 1 & 0 & \cdots 0 \end{vmatrix}}{\begin{vmatrix} u_1(0) & \cdots u_n(0) \\ \cdot\ \cdot & \cdots\ \cdot\ \cdot \\ u_1^{(n-1)}(0) & \cdots u_n^{(n-1)}(0) \end{vmatrix}}$$

$$= \frac{-\omega_i}{n\rho^{n-1}} e^{-\rho\omega_i y} P$$

---

* Milne, Bulletin of the American Mathematical Society, vol. 23 (1916–17), pp. 166–169.

where P is the quotient

$$\frac{\begin{vmatrix} P_0(\rho\omega_1)+\dfrac{E_{10}}{\rho^{m+1}}\cdots & P_0(\rho\omega_i)+\dfrac{E_{i0}}{\rho^{m+1}}\cdots & P_0(\rho\omega_n)+\dfrac{E_{n0}}{\rho^{m+1}} \\ \cdot\ \cdot\ \cdots & \cdot\ \cdot\ \cdots & \cdot\ \cdot \\ P_{n-2}(\rho\omega_1)+\dfrac{E_{1,n-2}}{\rho^{m+1}}\cdots & P_{n-2}(\rho\omega_i)+\dfrac{E_{i,n-2}}{\rho^{m+1}}\cdots & P_{n-2}(\rho\omega_n)+\dfrac{E_{n,n-2}}{\rho^{m+1}} \\ 0\ \cdot\ \cdot\ \cdots & 1\ \cdot\ \cdots & 0 \end{vmatrix}}{\begin{vmatrix} P_0(\rho\omega_1)+\dfrac{E_{10}}{\rho^{m+1}}\cdots & P_0(\rho\omega_n)+\dfrac{E_{n0}}{\rho^{m+1}} \\ \cdot\ \cdot\ \cdots & \cdot\ \cdot \\ P_{n-1}(\rho\omega_1)+\dfrac{E_{1,n-1}}{\rho^{m+1}}\cdots & P_{n-1}(\rho\omega_n)+\dfrac{E_{n,n-1}}{\rho^{m+1}} \end{vmatrix}_{y=0}}.$$

On expanding the determinants and performing the indicated division by the usual processes employed in dealing with asymptotic forms we find

$$v_i(y)=\frac{-\omega_i}{n\rho^{n-1}}e^{-\rho\omega_i y}\left\{\frac{\begin{vmatrix} P_0(\rho\omega_1) & \cdots P_0(\rho\omega_i) & \cdots P_0(\rho\omega_n) \\ \cdot & \cdot & \cdot \\ P_{n-2}(\rho\omega_1) & \cdots P_{n-2}(\rho\omega_i) & \cdots P_{n-2}(\rho\omega_n) \\ 0 & \cdots\ 1 & \cdots\ 0 \end{vmatrix}}{\begin{vmatrix} P_0(\rho\omega_1) & \cdots P_0(\rho\omega_n) \\ \cdot\ \cdot & \cdots\ \cdot\ \cdot \\ P_{n-1}(\rho\omega_1) & \cdots P_{n-1}(\rho\omega_n) \end{vmatrix}_{y=0}}+\frac{\overline{E}_i(y,\rho)}{\rho^{n+1}}\right\}$$

$$=\bar{v}_i(y)+\frac{-\omega_i}{n\rho^{n-1}}e^{-\rho\omega_i y}\frac{\overline{E}_i(y,\rho)}{\rho^{m+1}}.$$

Since $\omega_i\omega_j/\omega_1$ is an nth root of -1 reducing to ω_i for $j=1$, it is clear that $\bar{v}_i(y,\rho)\equiv\bar{v}_1(y,\rho\omega_i/\omega_1)$. For $\bar{v}_1(y)$ we have the development

$$\bar{v}_1(y,\rho)=\frac{-\omega_1}{n\rho^{n-1}}e^{-\rho\omega_1 y}\left(1+\sum_{l=1}^{l=m}\frac{B_l(y)}{(\rho\omega_1)^l}+\frac{A_1(y,\rho)}{\rho^{m+1}}\right)$$

where $A_1(y,\rho)$ is analytic at infinity in the ρ-plane, and the functions $B_l(y)$ are continuous together with their derivatives of all orders on $(0, 1)$ for $l=1, \cdots, m$. Thus

$$\bar{v}_i(y,\rho)=\frac{-\omega_i}{n\rho^{n-1}}e^{-\rho\omega_i y}\left(1+\sum_{l=1}^{l=m}\frac{B_l(y)}{(\rho\omega_i)^l}+\frac{A_i(y,\rho)}{\rho^{m+1}}\right),$$

$i=1, \cdots, n$, and

$$v_i(y,\rho)=\frac{-\omega_i}{n\rho^{n-1}}e^{-\rho\omega_i y}\left(1+\sum_{l=1}^{l=m}\frac{B_l(y)}{(\rho\omega_i)^l}+\frac{E_i(y,\rho)}{\rho^{m+1}}\right),$$

$i=1, \cdots, n$, as we were to show. Since the functions $B_l(y)$ are expressible in terms of the coefficients $A_{lk}(y)$ which occur in the asymptotic forms for $u_1^{(k)}, \cdots, u_n^{(k)}$ and which are independent of the sector S, they are independent of the sector S.

THEOREM XXXII. *If $f(x)$ is summable on $(0, 1)$; and if $G(x, y; \lambda)$ is the Green's function associated with the regular differential system of order $n=2\mu-1$ of the type described above; and if*

$$F_1{}^0{}_k(x,y,\rho) \equiv - \sum_{i=1}^{i=\mu} \sum_{s=0}^{s=k} (\rho\omega_i)^s \omega_i e^{\rho\omega_i(x-y)} \sum_{\alpha+\beta=k-s} A_{\alpha k}(x) B_\beta(y),$$

$$F_1{}^1{}_k(x,y,\rho) \equiv + \sum_{i=\mu+1}^{i=n} \sum_{s=0}^{s=k} (\rho\omega_i)^s \omega_i e^{\rho\omega_i(x-y)} \sum_{\alpha+\beta=k-s} A_{\alpha k}(x) B_\beta(y),$$

$$F_2{}^0{}_k(x,y,\rho) \equiv - \sum_{i=1}^{i=\mu-1} \sum_{s=0}^{s=k} (\rho\omega_i)^s \omega_i e^{\rho\omega_i(x-y)} \sum_{\alpha+\beta=k-s} A_{\alpha k}(x) B_\beta(y),$$

$$F_2{}^1{}_k(x,y,\rho) \equiv + \sum_{i=\mu}^{i=n} \sum_{s=0}^{s=k} (\rho\omega_i)^s \omega_i e^{\rho\omega_i(x-y)} \sum_{\alpha+\beta=k-s} A_{\alpha k}(x) B_\beta(y),$$

where $A_{\alpha k}(x)$ is the coefficient of $(\rho\omega_i)^{-\alpha}$ in the asymptotic form for $u_i^{(k)}(x)$ on S, $B_\beta(y)$ the coefficient of $(\rho\omega_i)^{-\beta}$ in the form for $v_i(y)$, then

$$\lim_{R\to\infty} \int_0^1 f(y) \int_{\gamma_1} \left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^{k+l} \left(n\rho^{n-1} \left\{ \frac{\partial^k G}{\partial x^k} ; \frac{\partial^k G}{\partial x^k} \right\} - \left\{ F_1{}^0{}_k ; F_1{}^1{}_k \right\} \right) d\rho dy = 0,$$

$$\lim_{R\to\infty} \int_0^1 f(y) \int_{\gamma_2} \left(1 - \frac{\rho^{4n}}{R^{4n}}\right)^{k+l} \left(n\rho^{n-1} \left\{ \frac{\partial^k G}{\partial x^k} ; \frac{\partial^k G}{\partial x^k} \right\} - \left\{ F_2{}^0{}_k ; F_2{}^1{}_k \right\} \right) d\rho dy = 0,$$

uniformly, $0<a\leqq x\leqq b<1$, $l\geqq 0$. *The expression*

$$\frac{1}{2\pi i} \frac{\partial^k}{\partial x^k} \int_0^1 f(y) \int_C \left(1 - \frac{\lambda^4}{\Lambda^4}\right)^{k+l} G(x,y;\lambda) d\lambda dy$$

is equivalent on any interval (a, b) completely interior to $(0, 1)$ to a linear combination with coefficients $A_{\alpha k}(x)$ of means of order $k+l$, $l\geqq 0$, formed from the Fourier series and their derived series to order k for the functions $f(x)B_0(x) \equiv f(x)$, $f(x)B_1(x), \cdots, f(x)B_k(x)$. On any interval (a, b) the problem of the derived series of Birkhoff of order $n=2\mu-1$ is reduced to a problem in the theory of derived Fourier series.

Case I. ρ on S_1'. We take $m=k$ in the asymptotic forms of Theorem II$'$ and Lemma XIV. We then compute the asymptotic form of

$$n\rho^{n-1}\left\{\frac{\partial^k}{\partial x^k}G(x,y\,;\,\rho^n)\,;\frac{\partial^k}{\partial x^k}\,G(x,y\,;\,\rho^n)\right\}$$

on S_1' by methods like those used in Theorem VII. We have

$$g(x,y)=\left\{\frac{1}{2}\sum_{i=1}^{i=n}u_i(x)v_i(y)\,;\,-\frac{1}{2}\sum_{i=1}^{i=n}u_i(x)v_i(y)\right\},$$

$$\left\{\frac{\partial^kG}{\partial x^k}\,;\frac{\partial^kG}{\partial x^k}\right\}=(-1)^n\frac{\begin{vmatrix} u_1^{(k)}(x) & \cdots & u_n^{(k)}(x) & \left\{\dfrac{\partial^k g}{\partial x^k}\,;\dfrac{\partial^k g}{\partial x^k}\right\} \\ W_1(u_1) & \cdots & W_1(u_n) & W_1(g) \\ \cdot & \cdots & \cdot & \cdot \\ W_n(u_1) & \cdots & W_n(u_n) & W_n(g)\end{vmatrix}}{\begin{vmatrix} W_1(u_1) & \cdots & W_1(u_n) \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ W_n(u_1) & \cdots & W_n(u_n)\end{vmatrix}}.$$

We multiply the first μ columns in the numerator by $\frac{1}{2}v_1(y),\cdots,\frac{1}{2}v_\mu(y)$ respectively, the next $\mu-1$ by $-\frac{1}{2}v_{\mu+1}(y),\cdots,\frac{1}{2}v_n(y)$, and add to the last. We then take the first term of the new last column outside the determinant. The result of substituting the asymptotic forms for $u_1,\cdots,u_n,v_1,\cdots,v_n$ is

$$n\rho^{n-1}\left\{\frac{\partial^kG}{\partial x^k}\,;\frac{\partial G^k}{\partial x^k}\right\}=\left\{F_1{}^0{}_k-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\,\frac{m_i(x,y,\rho)}{\rho}\,;\right.$$
$$\left.F_1{}^1{}_k+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\,\frac{m_i(x,y,\rho)}{\rho}\right\}+\frac{\Delta_1{}^{(k)}}{[\theta_0]+e^{\rho\omega_i}[\theta_1]},$$

where $\Delta_1{}^{(k)}$ is a determinant which differs from the determinant Δ_1 of Theorem VII only in having the elements of the first row changed to

$$\begin{aligned} a_{1j}&=(\rho\omega_j)^k e^{\rho\omega_j x}[1], & j&=1,\cdots,\mu\,;\\ a_{1j}&=(\rho\omega_j)^k e^{\rho\omega_j(x-1)}[1], & j&=\mu+1,\cdots,n\,;\\ a_{1,n+1}&=0. \end{aligned}$$

Thus we find

$$\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l}\left(n\rho^{n-1}\left\{\frac{\partial^kG}{\partial x^k}\,;\frac{\partial^kG}{\partial x^k}\right\}-\left\{F_1{}^0{}_k\,;\,F_1{}^1{}_k\right\}\right)=\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\,\frac{m_i}{\rho}\,;\right.$$
$$\left.+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\,\frac{m_i}{\rho}\right\}+\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^k\rho^k\left(\sum_{i=1}^{i=\mu}e^{\rho\omega_i x}m_{n+i}+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-1)}m_{n+i}\right)$$

where the functions m are uniformly bounded for all x and y on (0, 1) and for all ρ on S_1'. Consequently, by Lemmas III, IV, V,

$$\int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l}\left(n\rho^{n-1}\left\{\frac{\partial^k G}{\partial x^k};\frac{\partial^k G}{\partial x^k}\right\}-\left\{F_{1\,k}^{0};F_{1\,k}^{1}\right\}\right)d\rho$$

is uniformly bounded, $0<a\leqq x\leqq b<1$, $0\leqq y\leqq 1$, for all ρ on S_1'. Next, on integrating over (α, x) with respect to y just as we did in Theorem VIII, we find

$$\int_\alpha^x\int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l}\left(n\rho^{n-1}\left\{\frac{\partial^k G}{\partial x^k};\frac{\partial^k G}{\partial x^k}\right\}-\left\{F_{1\,k}^{0};F_{1\,k}^{1}\right\}\right)d\rho dy$$

$$=\int_\alpha^x\int_{\gamma_1}\left\{-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho};+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho}\right\}d\rho dy$$

$$+\int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{(k-1)+(l+1)}\rho^{k-1}\left(\sum_{i=1}^{i=\mu}e^{\rho\omega_i x}m_{n+i}+\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(x-1)}m_{n+i}\right)d\rho\,.$$

The expression on the right approaches zero uniformly, $0<a\leqq x\leqq b<1$, as $R\to\infty$ by Lemmas IV, V, VI. Thus

$$\lim_{R\to\infty}\int_\alpha^\rho\int_{\gamma_1}\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l}\left(n\rho^{n-1}\left\{\frac{\partial^k G}{\partial x^k};\frac{\partial^k G}{\partial x^k}\right\}-\left\{F_{1\,k}^{0};F_{1\,k}^{1}\right\}\right)d\rho dy=0,$$

uniformly, $0<a\leqq x\leqq b<1$.

With the aid of Theorem IV the present theorem is established in the case of the first integral.

Case II. ρ on S_2'. We compute the asymptotic form for

$$n\rho^{n-1}\left\{\frac{\partial^k G}{\partial x^k};\frac{\partial^k G}{\partial x^k}\right\}$$

in a manner analogous to that used in Case II of Theorm VII. We then follow our usual line of attack to validate the assertion concerning the second integral of this theorem.

We notice that in the case of Fourier series, as given in the differential system of order $n=2\mu-1$ analysed in §III,

$$\bar{F}_{1\,s}^{0}\equiv-\sum_{i=1}^{i=\mu}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)},$$

$$\bar{F}_{1\,s}^{1}\equiv+\sum_{i=\mu+1}^{i=n}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)},$$

$$\bar{F}_{2\,s}^{0}\equiv-\sum_{i=1}^{i=\mu-1}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)},$$

$$\bar{F}_{2\,s}^{1}\equiv+\sum_{i=\mu}^{i=n}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)}.$$

Thus

$$F_{j\,k}^{\,i} \equiv \sum_{s=0}^{s=k} \bar{F}_{j\,s}^{\,i} \sum_{\alpha+\beta=k-s} A_{\alpha k}(x) B_{\beta}(y), \qquad i = 0,1, \quad j = 1,2.$$

Since the functions A and B are independent of the sector S, the important relation between the derived series of Birkhoff and the derived series of Fourier, whose complete statement appears above, is now demonstrated.

It is now proposed to investigate a few of the leading properties of the derived series of Fourier series; some of them can be extended to Birkhoff series. The results which we shall present have all been obtained by W. H. Young,* who used methods different from ours.

THEOREM XXXIII. *If $\phi(x)$ is a k-fold integral in the sense of Lebesgue on $(0, 1)$, and if $G(x, y;\lambda)$ is the Green's function associated with the differential system $u'+\lambda u=0$, $u(0)-u(1)=0$, then*

$$\lim_{\nu\to\infty}\left(\frac{1}{2\pi i}\int_0^1 \phi^{(k)}(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\,\lambda)\,d\lambda\,dy\right.$$
$$\left. -\frac{1}{2\pi i}\,\frac{\partial^k}{\partial x^k}\int_0^1 \phi(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l} G(x,y\,;\,\lambda)\,d\lambda\,dy\right) = 0,$$

uniformly, $0<a\leqq x\leqq b<1$, for $l\geqq 0$. In other words, the Cesàro sums of order $k+l$, $l\geqq 0$, for the kth derived Fourier series of $\phi(x)$ and for the Fourier series of $\phi^{(k)}(x)$ are equivalent on (a, b).

By Theorems IX and XXXII we need consider merely the differences

$$\int_0^x\int_{\gamma_1}\left(1-\frac{\lambda^4}{\Lambda^4}\right)^{k+l}\left(\phi(y)\lambda^k(-1)^k e^{\lambda(y-x)}-\phi^{(k)}(y)e^{\lambda(y-x)}\right)d\lambda\,dy$$
$$+\int_1^x\int_{\gamma_2}\left(1-\frac{\lambda^4}{\Lambda^4}\right)^{k+l}\left(\phi(y)\lambda^k(-1)^k e^{\lambda(y-x)}-\phi^{(k)}(y)e^{\lambda(y-x)}\right)d\lambda\,dy.$$

By integration by parts we can show

$$\int_0^x \phi(y)(-1)^k\lambda^k e^{\lambda(y-x)}dy = \sum_{i=1}^{i=k}(-1)^{k-i+1}\lambda^{k-i}\left(\phi^{(i-1)}(x)-\phi^{(i-1)}(0)e^{-\lambda x}\right)$$
$$+\int_0^x \phi^{(k)}(y)e^{\lambda(y-x)}dy,$$

---

* W. H. Young, Proceedings of the London Mathematical Society, ser. 2, vol. 17 (1918), pp. 195–236.

$$\int_1^x \phi(y)(-1)^k\lambda^k e^{\lambda(y-x)}dy = \sum_{i=1}^{i=k}(-1)^{k-i+1}\lambda^{(k-i)}\left(\phi^{(i-1)}(x) - \phi^{(i-1)}(1)e^{\lambda(1-x)}\right)$$
$$+ \int_1^x \phi^{(k)}(y)e^{\lambda(y-x)}dy.$$

On substituting these expressions we find

$$\sum_{i=1}^{i=k}(-1)^{k-i+1}\phi^{(i-1)}(x)\int_C \lambda^{k-i}\left(1-\frac{\lambda^4}{\Lambda^4}\right)^{k+l}d\lambda$$
$$+\sum_{i=1}^{i=k}(-1)^{k-i}\phi^{(i-1)}(0)\int_{\gamma_1}\lambda^{k-i}\left(1-\frac{\lambda^4}{\Lambda^4}\right)^{(k-i)+(l+i)}e^{-\lambda x}d\lambda$$
$$+\sum_{i=1}^{i=k}(-1)^{k-i}\phi^{(i-1)}(1)\int_{\gamma_2}\lambda^{k-i}\left(1-\frac{\lambda^4}{\Lambda^4}\right)^{(k-i)+(l+i)}e^{\lambda(1-x)}d\lambda.$$

The integrals over C vanish identically by Cauchy's theorem. The integrals in the remaining terms approach zero uniformly, $0<a\leqq x\leqq b<1$ as $\Lambda\to\infty$, $l\geqq 0$, by the application of Lemma V.

By Theorem VI, Corollary I, the type of mean employed in summing the two series may be replaced by a Cesàro sum of the same order, $k+l$.

Corollary I. *The $(k+l)$th Cesàro sum of the kth derived Fourier series for $\phi(x)$ converges almost everywhere to $\phi^{(k)}(x)$; and converges uniformly on (a, b) if $\phi^{(k)}(x)$ is continuous.*

Corollary II. *We have, if $\psi(x)$ is bounded and summable on* (0,1),

$$\lim_{\nu\to\infty}\int_{x_1}^{x_2}\psi(\xi)\frac{1}{2\pi i}\frac{\partial^k}{\partial\xi^k}\int_0^1\phi(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}G(\xi,y\,;\,\lambda)d\lambda dy d\xi$$
$$=\int_{x_1}^{x_2}\psi(\xi)\phi^{(k)}(\xi)d\xi$$

for all x_1 and x_2 interior to (0, 1).

On applying the preceding theorem and its two corollaries to Theorem XXXII we find

Theorem XXXIV. *If $\phi(x)$ is a k-fold integral in the sense of Lebesgue, $0\leqq x\leqq 1$, then*

$$\lim_{\nu\to\infty}\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1\phi(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}G(x,y\,;\,\lambda)d\lambda dy = \phi^{(k)}(x)$$

almost everywhere, $0<x<1$; and if $\phi^{(k)}(x)$ is continuous, the convergence to the limit is uniform, $0<a\leqq x\leqq b<1$.

By Corollary I of Theorem XXXIII and by the last part of Theorem XXXII, we see that

$$\lim_{\nu\to\infty}\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1\phi(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}G(x,y\,;\lambda)d\lambda dy = \Phi_k(x)$$

almost everywhere, $0<x<1$, the convergence being uniform on (a, b) if $\phi^{(k)}(x)$ is continuous. By Corollary II of Theorem XXXIII we can compute the integral of $\Phi_k(x)$ by integrating under the limit sign on the left. We have

$$\Phi_{k-1}(x) - \Phi_{k-1}(x_1) = \int_{x_1}^{x}\Phi_k(\xi)d\xi,\ \Phi_k(x) = \Phi'_{k-1}(x) = \cdots = \Phi_0^{(k)}(x).$$

Since $\Phi_0(x)\equiv\phi(x)$, it follows that $\Phi_k(x)=\phi^{(k-1)}(x)$ and the theorem is proved.

THEOREM XXXV. *If $f(x)$ is summable on* $(0, 1)$, *then the convergence at a point x_0 interior to* $(0, 1)$, *of the $(k+l)$th Cesàro mean, $l\geqq 0$, of the kth derived series of the Fourier series for $f(x)$ is independent of the nature of the function outside an arbitrarily small neighborhood of the point x_0.**

We consider the differential system $u'+\lambda u=0$, $u(0)-u(1)=0$, starting from Theorem XXXII just as we did from Theorem XIX in proving Theorem XIV. The study of the $(k+l)$th mean of the kth derived Fourier series reduces to that of the expression

$$\frac{1}{2\pi}\int_0^{x_0}f(-z+x_0)\int_{-1}^{+1}(1-\phi^4)^{k+l}(-i)^k\phi^k\Lambda^{k+1}e^{-\Lambda i\phi z}d\phi dz$$
$$+\frac{1}{2\pi}\int_0^{1-x_0}f(z+x_0)\int_{-1}^{+1}(1-\phi^4)^{k+l}(-i)\phi^k\Lambda^{k+1}e^{\Lambda i\phi z}d\phi dz$$

where the integrals with respect to ϕ are taken over the real axis. We let A be an arbitrarily small fixed positive quantity, which we may suppose less than x_0 and $1-x_0$; we denote by h the function

$$\frac{d^{k+1}}{d\phi^{k+1}}\left[(1-\phi^4)^{k+l}\phi^k\right],$$

which is summable on $(-1, 1)$ for $l\geqq 0$. Then on integrating by parts $k+1$ times we find

$$\int_{-1}^{+1}(1-\phi^4)^{k+l}\phi^k\Lambda^{k+1}e^{\pm\Lambda i\phi z}d\phi = \frac{(\mp 1)^{k+1}}{z^{k+1}}\int_{-1}^{+1}he^{\pm\Lambda i\phi z}d\phi,$$

* W. H. Young, Proceedings of the London Mathematical Society, ser. 2, vol. 17 (1918) pp. 195–236.

an expression which is uniformly bounded for $z \geqq A$, $l \geqq 0$, and which approaches zero with $1/\Lambda$ for each positive value of z, by the theorem of Riemann-Lebesgue. It follows that in the expression above the integrals over (A, x_0) and $(A, 1-x_0)$ have the limit zero as $\Lambda \to \infty$. Therefore the convergence of the $(k+l)$th mean, $l \geqq 0$, at $x = x_0$ depends on the behavior of the sum

$$\frac{1}{2\pi}\int_0^A f(-z+x_0)\int_{-1}^{+1}(1-\phi^4)^{k+l}(-i)^k\phi^k\Lambda^{k+1}e^{-\Lambda i\phi z}d\phi dz$$
$$+\frac{1}{2\pi}\int_0^A f(z+x_0)\int_{-1}^{+1}(1-\phi^4)^{k+l}(-i)^k\phi^k\Lambda^{k+1}e^{\Lambda i\phi z}d\phi dz.$$

This establishes the theorem.

An immediate consequence of this result is

THEOREM XXXVI. *If $f(x)$ is summable on* $(0, 1)$, *the behavior of the expression*

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}G(x,y\,;\lambda)d\lambda dy, \quad l \geqq 0,$$

at the point $x = x_0$, $0 < x_0 < 1$, as $\nu \to \infty$ is independent of the nature of $f(x)$ outside an arbitrarily small neighborhood of x_0.

We now proceed to two other theorems of more general character, with a view to showing the effectiveness of the method here considered in discussing the derived series of Fourier series.

THEOREM XXXVII. *If $f(x)$ is summable on* $(0, 1)$; *and if*

$$F(x) = \frac{f(x)-f(2x_0-x)}{2(x-x_0)}$$

is summable on $(0, 1)$ *when x_0 is interior to* $(0, 1)$ *and $f(x)$ is defined for all values of x by the identity $f(x) \equiv f(x+1)$; and if $S_\Lambda^{(l)}$, $\Sigma_\Lambda^{(l)}$ denote the* $[\Lambda]$*th Cesàro sums of order l for a Fourier series and its first derived series respectively, then*

$$\lim_{\Lambda\to\infty}(\Sigma_\Lambda^{(l+1)}(f) - 5S_\Lambda^{(l+1)}(F) + 4S_\Lambda^{(l)}(F)) = 0, \quad x = x_0, \quad l \geqq 0.$$

In particular, if $f(x)$ is continuous and $\lim_{x\to x_0}F(x) = A$, *then*

$$\lim_{\Lambda\to\infty}\Sigma_\Lambda^{(l+1)}(f) = A, \quad x = x_0, \quad l > 0.*$$

---

* W. H. Young, Proceedings of the London Mathematical Society, ser. 2, vol. 17 (1918) pp. 195–236.

It is sufficient to demonstrate that

$$\lim_{\Lambda\to\infty}(\Sigma_\Lambda^{(1)}(f) - 5S_\Lambda^{(1)}(F) + 4S_\Lambda^{(0)}(F)) = 0, \quad x = x_0.$$

By taking $k=1$, $l=0$, in the final expression of Theorem XXXV and noticing that

$$\int_{-1}^{+1} (1-\phi^4)\phi\cos\Lambda\phi z d\phi = 0$$

we find that we can replace $\Sigma_\Lambda^{(1)}(f)$ by the expression

$$\frac{1}{2\pi}\int_0^A (f(z+x_0) - f(-z+x_0))\int_{-1}^{+1}(1-\phi^4)\phi\Lambda^2\sin\Lambda\phi z d\phi dz,$$

in the sense that this expression has the same behavior as $\Lambda\to\infty$. On integrating by parts we find

$$\int_{-1}^{+1}(1-\phi^4)\phi\Lambda^2\sin\Lambda\phi z d\phi = \frac{1}{z}\int_{-1}^{+1}(1-5\phi^4)\Lambda\cos\Lambda\phi z d\phi.$$

Hence we can replace $\Sigma_\Lambda^{(1)}(f)$ by

$$\frac{1}{2\pi}\int_0^A \frac{f(z+x_0) - f(-z+x_0)}{z}\int_{-1}^{+1}(1-5\phi^4)\Lambda\cos\Lambda\phi z d\phi dz.$$

We note that $F(-z+x_0)+F(z+x_0)=[f(z+x_0)-f(-z+x_0)]/z$. Thus we can replace $S_\Lambda^{(1)}(F)$ and $S_\Lambda^{(0)}(F)$ by

$$\frac{1}{2\pi}\int_0^A \frac{f(z+x_0) - f(-z+x_0)}{z}\int_{-1}^{+1}(1-\phi^4)\Lambda\cos\Lambda\phi z d\phi dz$$

and

$$\frac{1}{2\pi}\int_0^A \frac{f(z+x_0) - f(-z+x_0)}{z}\int_{-1}^{+1}\Lambda\cos\Lambda\phi z d\phi dz$$

respectively. When we make the three replacements in the expression

$$\Sigma_\Lambda^{(1)}(f) - 5S_\Lambda^{(1)}(F) + 4S_\Lambda^{(0)}(F)$$

we find that it then reduces to zero; that is to say, for $x=x_0$,

$$\lim_{\Lambda\to\infty}(\Sigma_\Lambda^{(1)}(f) - 5S_\Lambda^{(1)}(F) + 4S_\Lambda^{(0)}(F)) = 0.$$

The remainder of the theorem follows at once.

THEOREM XXXVIII. *The first derived series of the Fourier series for a function of bounded variation, $\chi(x)$, is summable (C, l), $l>0$, to $\chi'(x)$ almost everywhere on* $(0, 1)$.*

We employ here the concept of integration with respect to a function of bounded variation. We let x_0 be a point interior to $(0, 1)$ for which the function

$$X(z) = \frac{\chi(z + x_0) - \chi(-z + x_0)}{2} - z\chi'(x_0)$$

is defined and satisfies the conditions

$$X(+0) = 0, \qquad \lim_{z\to 0} \frac{1}{z}\int_0^z |dX| = 0.$$

Young has shown that the set of points thus defined is of measure 1.

Taking $G(x, y; \lambda)$ as the Green's function for the system $u'+\lambda u=0$, $u(0)-u(1)=0$, we find for $x=x_0$

$$\frac{\partial}{\partial x}\int_0^1 \chi(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l G(x,y;\lambda)d\lambda dy$$

$$= \int_0^{x_0} \chi(y)\int_{\gamma_1}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l \lambda e^{\lambda(y-x_0)}d\lambda dy$$

$$+ \int_{x_0}^1 \chi(y)\int_{\gamma_2}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l \lambda e^{\lambda(y-x_0)}d\lambda dy$$

$$+ \int_0^1 \chi(y)\int_{\gamma_1}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l \frac{-\lambda e^{-\lambda x_0}e^{\lambda(y-1)}}{1-e^{-\lambda}}d\lambda dy$$

$$+ \int_0^1 \chi(y)\int_{\gamma_2}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^l \frac{\lambda e^{\lambda(1-x_0)}e^{\lambda y}}{e^{\lambda}-1}d\lambda dy.$$

In the third integral we have

$$\left|\int_0^1 \chi(y)\lambda e^{\lambda(y-1)}dy\right| = \left|\chi(1) - \chi(0)e^{-\lambda} - \int_0^1 e^{\lambda(y-1)}d\chi\right|$$

$$\leqq |\chi(1)| + |\chi(0)| + \int_0^1 |d\chi| \leqq M,$$

---

* W. H. Young, Proceedings of the London Mathematical Society, ser. 2, vol. 13 (1914) pp. 13–28.

so that for $l>0$ it approaches zero with $1/\Lambda$ in accord with Lemma V. By a similar argument we can dispose of the fourth integral. The two remaining integrals can be changed into the form

$$\frac{1}{2\pi}\int_0^{x_0}\chi(-z+x_0)\int_{-1}^{+1}(1-\phi^4)^l(-i\phi)\Lambda^2 e^{-\Lambda i\phi z}d\phi dz$$
$$+\frac{1}{2\pi}\int_0^{1-x_0}\chi(z+x_0)\int_{-1}^{+1}(1-\phi^4)^l(-i\phi)\Lambda^2 e^{\Lambda i\phi z}d\phi dz,$$

as was done in Theorems XIV and XXXV. We take A positive but less than x_0 and $1-x_0$. Then we show the integrals over (A, x_0) and $(A, 1-x_0)$ negligible, in the following manner: we integrate by parts with respect to z in each of them and apply the inequalities of Lemma VII to the resulting expressions. The discussion has many points of similarity with the first part of that given in Theorem XIV. Thus we consider

$$\frac{1}{\pi}\int_0^A (X(z)+z\chi'(x_0))\int_{-1}^{+1}(1-\phi^4)\phi\Lambda^2\sin\Lambda\phi z d\phi dz, \quad l>0.$$

It is readily shown that

$$\lim_{\Lambda\to\infty}\frac{1}{\pi}\int_0^A z\int_{-1}^{+1}(1-\phi^4)^l\phi\Lambda^2\sin\Lambda\phi z d\phi dz = 1.$$

To prove the present theorem, then, it remains to show that

$$\lim_{\Lambda\to\infty}\int_0^A X(z)\int_{-1}^{+1}(1-\phi^4)^l\phi\Lambda^2\sin\Lambda\phi z d\phi dz = 0.$$

This is easily accomplished after an integration by parts with the aid of the inequalities of Lemma VII; the method is that used by Hardy for the proof of the last step in Theorem XIV and by Young for the present theorem. This finishes the proof.

We leave the reader to apply the last two theorems to the study of the derived series of Birkhoff of order $n=2\mu-1$.

VII. The series for $n=2\mu$

The regular differential system

$$u^{(n)}+\bigstar+p_2u^{(n-2)}+\cdots+(p_n+\lambda)u=0,$$
$$W_1(u)=0,\cdots,W_n(u)=0, \qquad n=2\mu,$$

where

$$W_i(u)\equiv\alpha_i u^{(k_i)}(0)+\beta_i u^{(k_i)}(1)+\cdots,$$
$$n-1\geqq k_1\geqq k_2\geqq\cdots\geqq k_n, \quad k_{i+2}>k_i,$$

presents certain salient differences from the case $n=2\mu-1$; but these are differences of form which do not necessitate any essential alteration in the spirit of the method used.

We shall write $\lambda=\rho^n$, considering instead of the whole λ-plane two adjacent sectors S: $\mathcal{R}(\rho\omega_1)\leqq\mathcal{R}(\rho\omega_2)\leqq\cdots\leqq\mathcal{R}(\rho\omega_n)$ in the ρ-plane. This region we shall call Σ. The sector S is a region T of §I for $c=0$. On employing the asymptotic forms for solutions $u_1,\cdots,u_n$ on a region T including S in its interior we ascertain the distribution of the characteristic values in S.* On one of the bounding rays of a sector S, $\mathcal{R}(\rho\omega_\mu)=\mathcal{R}(\rho\omega_{\mu+1})=0$; we arrange the two adjacent sectors S forming Σ so that they have such a bounding ray in common. Then the characteristic values of ρ on Σ are distributed in two series, asymptotically near points equally spaced at distance 2π along each of two lines parallel to the bisecting ray of Σ; these are simple characteristic values except in the case of a value common to the two series, which is always double.

We now prepare the regions Σ', S' in a manner entirely analogous to that employed in the case $n=2\mu-1$. The arcs Γ, γ and the circles C can then be defined. The notations σ, R will be carried over from our preceding work.

On any region S we have†

$$\mathcal{R}(\rho\omega_1)\leqq\mathcal{R}(\rho\omega_2)\leqq\cdots\leqq\mathcal{R}(\rho\omega_{\mu-1})\leqq-\beta<0\,;$$
$$\mathcal{R}(\rho\omega_\mu)=\mathcal{R}(\rho\omega_{\mu+1})=0,\text{ along one bounding ray;}$$
$$\omega_\mu=-\omega_{\mu+1}\,;$$
$$\mathcal{R}(\rho\omega_n)\geqq\mathcal{R}(\rho\omega_{n-1})\geqq\cdots\geqq\mathcal{R}(\rho\omega_{n-2})\geqq\beta>0.$$

Hence we can consider the whole sector S' at once; but we have two roots of -1, namely ω_μ, $\omega_{\mu+1}$, which have to receive special attention. This is the reverse of the situation in the case $n=2\mu-1$, where the sector S' had to be divided and where there was on each sector S_1', S_2' a single root of -1 which required separate consideration.

In the present discussion we have to replace Lemmas IV, V, VI. We have

Lemma IV′. *If $m(x,\rho)$ is a function of x and ρ bounded for $0<a\leqq x\leqq b<1$ and for all ρ on S', then*

$$\lim_{R\to\infty}\int_\gamma \rho^k e^{\rho\omega_i x} m\,d\rho=0 \qquad (i=1,\cdots,\mu-1),$$

* Birkhoff, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), p. 117.

† Birkhoff, these Transactions, vol. 9 (1908), p. 386.

$$\lim_{R\to\infty}\int_\gamma \rho^k e^{\rho\omega_i(x-1)} m d\rho = 0 \qquad (i = \mu+2, \cdots, n),$$

for any k, uniformly, $0<a\leqq x\leqq b<1$.

LEMMA V'. *If $m(x, y, \rho)$ is a function of x, y, and ρ bounded for $0<a\leqq x \leqq b<1$, $0\leqq y\leqq 1$, and for all ρ on S', then*

$$\int_\gamma \rho^k\left(1-\frac{\rho^{4n}}{R^{4n}}\right) e^{\rho\omega_\mu x} m d\rho,$$

$$\int_\gamma \rho^k\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l} e^{\rho\omega_{\mu+1}(x-1)} m d\rho,$$

for any $k\geqq 0$, (1) are uniformly bounded, $0<a\leqq x\leqq b<1$, $0\leqq y\leqq 1$, for all γ on S' when $l=0$, and (2) approach zero uniformly with $1/R$ for $0<a\leqq x\leqq b <1$, $0\leqq y\leqq 1$, when $l>0$.

LEMMA VI'. *If $m(x, y, \rho)$ is uniformly bounded, $0\leqq x\leqq 1$, $0\leqq y\leqq 1$, for all ρ on S', then*

$$\lim_{R\to\infty}\int_\alpha^x\int_\gamma e^{\rho\omega_i(x-y)}\frac{m}{\rho}\,d\rho dy = 0 \qquad (i = 1, \cdots, \mu),$$

uniformly, $0\leqq \alpha\leqq x\leqq 1$; and

$$\lim_{R\to\infty}\int_\alpha^x\int_\gamma e^{\rho\omega_i(x-y)}\frac{m}{\rho}\,d\rho dy = 0 \qquad (i = \mu+1, \cdots, n),$$

uniformly, $0\leqq x\leqq \alpha\leqq 1$.

The proofs are entirely analogous to those used for the preceding lemmas. Lemma III does not need to be replaced.

The Green's function for a system of order $n=2\mu$ is given by the explicit formula written down in Theorem VII. We proceed to put this expression in a more useful form by multiplying the first μ columns in the numerator by $\frac{1}{2}v_1(y), \cdots, \frac{1}{2}v_\mu(y)$ respectively, the next μ by $-\frac{1}{2}v_{\mu+1}(y), \cdots, -\frac{1}{2}v_n(y)$ respectively, and adding to the last. We then take the first term in the new last column outside the determinant. The result of substituting the asymptotic forms of $u_1, \cdots, u_n, v_1, \cdots, v_n$ is then*

$$n\rho^{n-1}G = \left\{-\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}[\omega_i]\,;+\sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}[\omega_i]\right\} + \frac{\Delta_3}{[\theta_1]e^{2\rho\omega_\mu}+[\theta_0]e^{\rho\omega_\mu}+[\theta_2]}$$

* Tamarkin, Rendiconti del Circolo Matematico di Palermo, vol. 34 (1912), p. 364.

where Δ_3 is a determinant of order $n+1$ in which the element a_{hj}, $h=0, 1, \cdots, n$, $j=1, 2, \cdots, n+1$ may described as follows:

$$a_{0,j}=e^{\rho\omega_j x}[1], j=1, \cdots, \mu\,; \; a_{0,j}=e^{\rho\omega_j(x-1)}[1], j=\mu+1, \cdots, n\,; \; a_{0,n+1}=0\,;$$

and for $h=1, 2, \cdots, n$,

$$a_{hj} = [\alpha_h\omega_j^{k_h}], j = 1, \cdots, \mu-1\,; \; a_{h,\mu} = [\alpha_h\omega_\mu^{k_h}] + e^{\rho\omega_\mu}[\beta_h\omega_\mu^{k_h}]\,;$$

$$a_{h,\mu+1} = [\alpha_h\omega_{\mu+1}^{k_h}]\, e^{-\rho\omega_{\mu+1}} + [\beta_h\omega_{\mu+1}^{k_h}]\,; \; a_{hj} = [\beta_h\omega_j^{k_h}],\; j = \mu+2, \cdots, n\,;$$

$$a_{h,n+1} = -\sum_{i=1}^{i=\mu} e^{\rho\omega_i(1-y)}[\beta_h\omega_i^{k_h+1}] + \sum_{i=\mu+1}^{i=n} e^{-\rho\omega_i y}[\alpha_h\omega_i^{k_h+1}].$$

On S', $1/\{[\theta_1]e^{2\rho\omega_\mu}+[\theta_0]e^{\rho\omega_\mu}+[\theta_2]\}$ is bounded.* In short, we are prepared to proceed as we did in the case $n=2\mu-1$.

We can thus establish Theorems IX′, XII′, XIII′, whose statements are identical with those of Theorems IX, XII, XIII, respectively, except for the change from $n=2\mu-1$ to $n=2\mu$. In the proof of Theorem XII′, essentially the same as that of Theorem XII, we must remark a slight difference. If we let $[\sigma]$ be that part of σ on S', then

$$\int_0^1 \int_{[\sigma]} f(y) n\rho^{n-1} G(x, y\,;\rho^n)\, d\rho\, dy$$

approaches zero uniformly, $0 \leqq x \leqq 1$, as the center of σ recedes indefinitely from the origin. Hence on Σ,

$$\int_0^1 \int_{\sigma} f(y) n\rho^{n-1} G(x, y\,;\rho^n)\, d\rho\, dy$$

approaches zero in similar fashion. Because of the distribution of the characteristic values when $n=2\mu$ the circles σ remote from the origin may be intersected by the bisecting ray of Σ, and it is necessary to introduce the corresponding modification of the proof. Theorem XIII′, which asserts the equivalence of Birkhoff series of order $n=2\mu$ and Fourier series on any interval (a, b) completely interior to $(0, 1)$, is the strongest theorem possible, in general, as we see by reference to the sine, cosine, and Fourier series in the case $n=2$.

We next introduce

LEMMA VIII′. *On any region S'*

$$\frac{[\theta_1]e^{2\rho\omega_\mu} + [\theta_0]e^{\rho\omega_\mu} + [\theta_2]}{[\bar\theta_1]e^{2\rho\omega_\mu} + [\bar\theta_0]e^{\rho\omega_\mu} + [\bar\theta_2]} = [1]$$

if $\theta_0 = \bar\theta_0, \theta_1 = \bar\theta_1, \theta_2 = \bar\theta_2$.

---

* Birkhoff, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), p. 120.

Lemma IX′. *The integrals*

$$\int_\gamma \int_\alpha^x e^{\rho\omega_i(1-y)}(m/\rho)\, d\rho dy \qquad (i = 1, \cdots, \mu),$$

$$\int_\gamma \int_\alpha^x e^{-\rho\omega_i y}(m/\rho)\, d\rho dy \qquad (i = \mu+1, \cdots, n),$$

where $m(x, y, \rho)$ is bounded for all x and y on (0, 1) *and for all ρ on S', converge uniformly to zero with* $1/R$, $0 \leqq x \leqq 1$, $0 \leqq \alpha \leqq 1$.

We can now obtain Theorems XVII′ and XVIII′ which may be stated merely by replacing $n = 2\mu - 1$ by $n = 2\mu$ in Theorems XVII and XVIII. It is to be noticed that Theorem XVIII′ includes as special cases the leading results of Haar and of Walsh cited in the introduction.

We next introduce the determinant $\mathcal{D}(x, y, \rho)$ obtained from Δ_3 by replacing each asymptotic expression $[A]$ by its dominating constant term A, and deleting the terms involving $e^{\rho\omega_\mu}$, $e^{\rho\omega_{\mu+1}}$. From $\mathcal{D}(0, y, \rho)$ and $\mathcal{D}(1, y, \rho)$ we obtain new determinants $D^0(y, \rho)$, $D^1(y, \rho)$ by replacing by zero the terms of the first rows explicitly involving the variable ρ. We then obtain two theorems similar to Theorems XIX and XX.

Theorem XIX′. *If $f(x)$ is summable on* (0, 1), *then*

$$\lim_{R\to\infty} \int_0^1 f(y) \int_\gamma \Bigg(n\rho^{n-1}G(x, y\,;\rho^n) - \Bigg\{ -\sum_{i=1}^{i=\mu} e^{\rho\omega_i(x-y)}\omega_i\,;\, + \sum_{i=\mu+1}^{i=n} e^{\rho\omega_i(x-y)}\omega_i \Bigg\} - \frac{1}{\theta_2}\mathcal{D}(x, y, \rho)\Bigg) d\rho dy = 0$$

uniformly, $0 \leqq x \leqq 1$.

Theorem XX′. *If $f(x)$ is summable on* (0,1) *then*

$$\lim_{R\to\infty} \int_0^1 f(y) \int_\gamma \Bigg(n\rho^{n-1}G(0, y\,;\rho^n) - \sum_{i=\mu+1}^{i=n} \omega_i e^{-\rho\omega_i y} \frac{1}{\theta_2} D^0(y, \rho) \Bigg) d\rho dy = 0,$$

$$\lim_{R\to\infty} \int_0^1 f(y) \int_\gamma \Bigg(n\rho^{n-1}G(1, y\,;\rho^n) + \sum_{i=1}^{i=\mu} \omega_i e^{\rho\omega_i(1-y)} \frac{1}{\theta_2} D^1(y, \rho) \Bigg) d\rho dy = 0.$$

We can next use Lemmas X and XII much as we did in §IV. In fact, we have, with the notations of Lemma X and Theorem XXI,

$$\int_0^1 f(y) \int_\gamma \omega_k e^{\rho\omega_k(1-y)} d\rho dy = \int_0^1 f(y)\, \frac{e^{-R\lambda_1 y} - e^{-R\lambda_2 y}}{y}\, dy$$

where $\theta_2-\theta_1=-\pi/n$, for $k=1,\cdots,n$; and

$$\int_0^1 f(y)\int_\gamma \omega_k e^{-\rho\omega_k y}d\rho dy = \int_0^1 f(y)\frac{e^{-R\lambda_1 y}-e^{-R\lambda_2 y}}{y}dy$$

where $\theta_2-\theta_1=\pi/n$, for $k=\mu+1,\cdots,n$. Hence we obtain Theorem XXI′, whose statement parallels that of Theorem XXI. Likewise there results

THEOREM XXII′. *If $\phi(x)$ is summable on* (0, 1), *and if*

$$\Phi_1(x)=\frac{1}{x}\int_0^x\phi(x)dx,\qquad \Phi_2(x)=\frac{1}{x}\int_0^x\phi(1-x)dx$$

are of bounded variation near $x=0$, then

$$\lim_{R\to\infty}\frac{1}{2\pi i}\int_0^1\phi(y)\int_\gamma n\rho^{n-1}G(k,y\,;\rho^n)d\rho dy = A_k\Phi_1(+0)+B_k\Phi_2(+0)$$

where A_k, B_k are constants depending only upon the boundary conditions, $k=0,1$.

For specific formulas concerning A_k, B_k, we refer to a paper of Birkhoff.*

In order to bring the discussion of the adjoint series of order $n=2\mu$ under that already carried out, we employ the device of writing $\bar{x}=1-x$, $\bar{y}=1-y$, $\bar{f}(\bar{y})=f(1-y)$, whereby

$$\int_0^1 f(y)\int_C G(y,x\,;\lambda)d\lambda dy=\int_0^1\bar{f}(\bar{y})\int_C G(1-\bar{y},1-\bar{x}\,;\lambda)d\lambda dy.$$

Then on S' we find for $n\rho^{n-1}G(1-\bar{y},1-\bar{x};\rho^n)$ an asymptotic form which in its essentials resembles that for $n\rho^nG(\bar{x},\bar{y};\rho^n)$. We are then able to prove Theorems XXIII′, XXIV′, XXV′, XXVI′, XXVII′, differing from the corresponding theorems of §V only in having $n=2\mu-1$ replaced by $n=2\mu$.

The theorems on convergence at the end points $x=0$, $x=1$ present certain differences. We define $\mathfrak{D}^0$ and $\mathfrak{D}^1$ from $\mathcal{D}(y,0,\rho)$ and $\mathcal{D}(y,1,\rho)$ by replacing by zero certain terms of the last columns, as in the corresponding situation in §V. Then we have

THEOREM XXVIII′. *If $f(x)$ is summable on* (0, 1), *then*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_\gamma\Big(n\rho^{n-1}G(y,x;\rho^n)-\Big\{\sum_{i=\mu+1}^{i=n}e^{\rho\omega_i(y-x)}\omega_i\,;\,-\sum_{i=1}^{i=\mu}e^{\rho\omega_i(y-x)}\omega_i\Big\}$$
$$-\frac{1}{\theta_2}\mathcal{D}(y,x,\rho)\Big)d\rho dy=0$$

uniformly, $0\leqq x\leqq 1$.

---

* Birkhoff, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), pp. 125–26.

THEOREM XXIX′. *If $f(x)$ is summable on* (0, 1), *then the limit in Theorem* XXVIII′ *is valid if x is set equal to* 0 *or* 1 *and $\mathfrak{D}$ is replaced by the corresponding determinant $\mathfrak{D}$; for example,*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_\gamma\Big(n\rho^{n-1}G(y,0\,;\,\rho^n)+\sum_{i=1}^{i=\mu}e^{\rho\omega_i y}\omega_i-\frac{1}{\theta_2}\mathfrak{D}^0(y,\rho)\Big)d\rho dy=0.$$

Theorem XXX′ is the exact analogue of Theorem XXX.

THEOREM XXXI′. *If $\phi(x)$ is summable on* (0, 1), *and if*

$$\Phi_1(x)=\frac{1}{x}\int_0^x\phi(x)dx,\quad \Phi_2(x)=\frac{1}{x}\int_0^x\phi(1-x)dx$$

are of bounded variation near $x=0$, then

$$\lim_{R\to\infty}\frac{1}{2\pi i}\int_0^1\phi(y)\int_\gamma n\rho^{n-1}G(y,k\,;\,\rho^n)d\rho dy=\mathfrak{A}_k\Phi_1(+0)+\mathfrak{B}_k\Phi_2(+0)$$

where $\mathfrak{A}_k$ and $\mathfrak{B}_k$ are constants depending only upon the boundary conditions, $k=0, 1$. *For example,*

$$\mathfrak{A}_0\Phi_1(+0)+\mathfrak{B}_0\Phi_2(+0)\equiv\frac{\mu}{2n}\Phi_2(+0)$$

$$+\frac{1}{2n\theta_2}\begin{vmatrix} a_1 & \cdots & a_\mu & a_{\mu+1} & \cdots & a_n & 0\\ \alpha_1\omega_1^{k_1} & \cdots & \alpha_1\omega_\mu^{k_1} & \beta_1\omega_{\mu+1}^{k_1} & \cdots & \beta_1\omega_n^{k_1} & -\beta_1\sum_{i=1}^{i=\mu}\omega_i^{k_1+1}\\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot\\ \alpha_n\omega_1^{k_n} & \cdots & \alpha_n\omega_\mu^{k_n} & \beta_n\omega_{\mu+1}^{k_n} & \cdots & \beta_n\omega_n^{k_n} & -\beta_n\sum_{i=1}^{i=\mu}\omega_i^{k_n+1}\end{vmatrix},$$

where $a_i=-\Phi_1(+0)/\omega_i$, $i=1, 2, \cdots, \mu$, and $a_i=\Phi_2(+0)/\omega_i$, $i=\mu+1, \cdots, n$.

Lastly, we pass to the consideration of the term-by-term derived series,

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}G(x,y\,;\,\lambda)d\lambda dy\qquad (k=1,2,\cdots),$$

under the assumption that the coefficients $p_2, \cdots, p_n$ of the differential equation are continuous together with their derivatives of all orders. Our theorems deal with the method of summing the series given by the expression

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\Big(1-\frac{\lambda^s}{\Lambda_\nu^s}\Big)^{k+l}G(x,y\,;\,\lambda)d\lambda dy,\qquad l\geqq 0.$$

Lemmas XIII and XIV hold for $n=2\mu$ as well as for $n=2\mu-1$. If we write

$$F_k^0(x,y,\rho) \equiv -\sum_{i=1}^{i=\mu}\sum_{s=0}^{s=k}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)}\sum_{\alpha+\beta=k-s}A_{\alpha k}(x)B_\beta(y),$$

$$F_k^1(x,y,\rho) \equiv +\sum_{i=\mu+1}^{i=n}\sum_{s=0}^{s=k}(\rho\omega_i)^s\omega_i e^{\rho\omega_i(x-y)}\sum_{\alpha+\beta=k-s}A_{\alpha k}(x)B_\beta(y),$$

we find

$$n\rho^{n-1}\left\{\frac{\partial^k}{\partial x^k}G\,;\frac{\partial^k}{\partial x^k}G\right\} \equiv \left\{F_k^0 - \sum_{i=1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho}\,;F_k^1+\sum_{i=\mu+1}^{i=\mu}e^{\rho\omega_i(x-y)}\frac{m_i}{\rho}\right\}$$
$$+\frac{\Delta_3^{(k)}}{[\theta_1]e^{2\rho\omega_\mu}+[\theta_0]e^{\rho\omega_\mu}+[\theta_2]}$$

where $\Delta_3^{(k)}$ is a determinant differing from Δ_3 only in having the elements of the first row given by

$$a_{0j} = (\rho\omega_j)^k e^{\rho\omega_j x}[1] \qquad (j=1,\cdots,\mu);$$
$$a_{0j} = (\rho\omega_i)^k e^{\rho\omega_j(x-1)}[1] \qquad (j=\mu+1,\cdots,n);$$
$$a_{0,n+1} = 0,$$

for ρ on S'. The demonstration of Theorem XXXII′ then follows the lines of that of Theorem XXXII.

Theorem XXXII′. *If $f(x)$ is summable on $(0,1)$and if $G\,(x,y;\lambda)$ is the Green's function associated with the regular differential system of order $n=2\mu$ of the type described above; and if $F_k^0(x,y,\rho)$, $F_k^1(x,y,\rho)$ are defined by the identities above, then*

$$\lim_{R\to\infty}\int_0^1 f(y)\int_\gamma\left(1-\frac{\rho^{4n}}{R^{4n}}\right)^{k+l}\left(n\rho^{n-1}\left\{\frac{\partial^k G}{\partial x^k}\,;\frac{\partial^k G}{\partial x^k}\right\}-\left\{F_k^0\,;F_k^1\right\}\right)d\rho\,dy = 0$$

uniformly, $0<a\leqq x\leqq b<1$, for $l\geqq 0$. The expression

$$\frac{1}{2\pi i}\frac{\partial^k}{\partial x^k}\int_0^1 f(y)\int_{C_\nu}\left(1-\frac{\lambda^4}{\Lambda_\nu^4}\right)^{k+l}G(x,y\,;\lambda)d\lambda\,dy$$

is therefore equivalent on any interval (a,b) completely interior to $(0,1)$ to a linear combination with coefficients $A_{\alpha k}(x)$ of means of order $k+l$, $l\geqq 0$, formed from the Fourier series and their derived series up to order k for the functions $f(x)B_0(x)\equiv f(x)$, $f(x)B_1(x)$, $\cdots$, $f(x)B_k(x)$. On any interval (a,b) the problem of derived Birkhoff series of order $n=2\mu$ is reduced to a problem in Fourier series and their derived series.

Theorems XXXIV′ and XXXVI′, analogous to the correspondingly numbered theorems of §VI, can then be obtained without difficulty.

From our brief indications the reader will perceive that the case $n = 2\mu$ is in its main features similar to the case $n = 2\mu - 1$, which we have discussed in greater detail.

Columbia University,
New York, N.Y.

ANALYTIC APPROXIMATIONS TO TOPOLOGICAL TRANSFORMATIONS*

BY

PHILIP FRANKLIN AND NORBERT WIENER

I. Introduction

A continuous transformation of a two-dimensional region into itself, or into a second such region may be characterized by a pair of continuous functions

$$X = X(x, y), \quad Y = Y(x, y).$$

The continuity of these functions, and the existence of a unique pair of inverse functions are the only restrictions imposed by the topologist. On the other hand, the differential geometer usually demands of the transformations he considers, that the functions defining them be analytic. It is a matter of some interest to investigate the relations connecting these two types of transformation, and in particular to determine whether the general transformations of the first type may not be regarded as in some sense limiting cases of the more special analytic type. This question is here discussed, and we shall prove that *every continuous one-to-one transformation of a two-dimensional region of finite connectivity may be approximated to an arbitrary degree of exactness by an analytic one-to-one transformation.*†

The analogy between the present investigation and the Weierstrass approximation theorem is obvious. The peculiar difficulty of the problem here treated is the need of keeping our approximating transformation one-to-one. This requires a discussion of a combinatorial nature.

II. Continuous transformations

Let the transformation

$$X = X(x, y), \quad Y = Y(x, y)$$

transform the closed region r of the x, y plane into the closed region R of the X, Y plane. Let $X(x, y)$ and $Y(x, y)$ be continuous and single-valued, and let it be possible to write

$$x = x(X, Y), \quad y = y(X, Y),$$

---

* Presented to the Society, May 2, 1925; received by the editors in May, 1925, and (revised) in March, 1926.

† This theorem was suggested to the authors by Professor J. W. Alexander.

where these functions are likewise single-valued. We shall then call our transformation of r into R *biunivocal* (or *one-to-one*) and *continuous*. We shall show that the inverse of this transformation is likewise continuous.

The continuity of this inverse transformation from R to r will follow if we show that any sequence of points $P_1, P_2, \cdots$ approaching a limit P is transformed into a sequence $p_1, p_2, \cdots$ approaching a limit p. We form the set p_i, the transform of the set P_i, and notice that, since it is an infinite set in the closed region r, it must contain at least one subsequence with a limit point, say p. That is,

$$\lim_{n\to\infty} p_{i_n} = p.$$

From the continuity and biunivocal character of the direct transformation, it follows that

$$\lim_{n\to\infty} P_{i_n} = \lim_{n\to\infty} P_n = P.$$

Since this argument proves that any subset of the p_i approaching a limit must approach p, we conclude that the set p_i approaches p as a limit.

In order to discuss approximations to transformations, we shall need a measure of the *distance* between two transformations. If S and T are two transformations, with inverses S^{-1} and T^{-1} respectively, then we define the distance between them as the greater of the two numbers

$$\text{maximum } [S(P)-T(P)] \text{ and maximum } [S^{-1}(P)-T^{-1}(P)],$$

where $[A-B]$ denotes the distance from A to B, and the maximum is taken as P ranges over the entire region for which the transformations are defined. To approximate to within a given distance of a transformation T, it is sufficient to approximate for the direct transformation alone. This follows from the fact that if S_n are a series of transformations such that

$$\lim_{n\to\infty} \text{maximum } [S_n(P) - T(P)] = 0,$$

then

$$\lim_{n\to\infty} \text{maximum } [S_n^{-1}(P) - T^{-1}(P)] = 0.$$

For, we have

$$[S_n^{-1}(P) - T^{-1}(P)] = [T^{-1}TS_n^{-1}(P) - T^{-1}S_nS_n^{-1}(P)],$$

and hence, since S_n is approaching T, and T^{-1} is continuous, our contention follows.

As we shall wish our approximating transformations to extend over the entire regions of definition of the transformation and its inverse, r and R respectively, if we kept to these regions we should have to take the boundary of r into that of R. In general this can not be done by any analytic transformation. We avoid this difficulty by seeking a transformation which shall exist in some region including r, whose inverse exists in some region including R, and which approximates the given transformation in the sense of having its distance from it, as defined above, arbitrarily small, say less than ϵ.

We restrict ourselves to the case where the region r, and hence R, is of finite connectivity, bounded by a number of simple, closed Jordan curves. We embed r and R in two large squares, s and S respectively, so chosen that no point of r is within a distance ϵ of s, and no point of R is within a distance ϵ of S. We then map the region $S-R$ on the region $s-r$ by a continuous transformation, which agrees with the original transformation on the boundary of R and r, and takes the vertices of the first square into those of the second. This is seen to be possible by linear interpolation in the case where r and R are bounded by circles, and the general case reduces to this in virtue of the Jordan-Schoenflies theorem.

If, now, we obtain an analytic biunivocal transformation T_2 which exists in the square s and whose inverse exists in a square inside of S, concentric with it and at distance ϵ from it such that the distance between the transformations T_2 and T_1, the extended continuous transformation, is less than ϵ, we shall have solved our problem for the original regions r and R.

III. The polygonal network

In approximating the extended transformation, T_1, which takes the interior, sides and vertices of a square s into those of a square S, we shall find a simultaneous subdivision of s and S of great service. In this section we shall prove

Lemma 1. *Given a continuous, biunivocal transformation T_1 which takes the interior, sides, and vertices of a square s into those of a square S, we may subdivide s and S into corresponding convex polygons, of which the number meeting at a vertex is one, two, or three according as the vertex is a vertex of the square, a point on a side of the square, or an interior point of the square, respectively, and having the property that any continuous biunivocal transformation T_2 which maps each polygon of s on the corresponding polygon of S is at distance less than $\epsilon/2$ from T_1.*

Let H be selected less than 10ϵ, and $\eta(H)$ be chosen less than H, and so that

$$[P_1 - P_2] < H \text{ if } [p_1 - p_2] < \eta.$$

Let us place on s a square network with mesh $\eta(H)$, and sides parallel to the sides of s, and find its transform under T_1. Let $N_1, N_2, \cdots$ be the intersection points of this transformed network in S, including the points where the network cuts the boundary. Let the least distance between any such point and a point on a side of the net not abutting on it be $3H'$. Surround each of the points $N_1, N_2, \cdots$ in S with a circle with radius H'. No two such circles can intersect. Furthermore, no such circle about N_k can intersect a side of a compartment of the net which does not terminate in N_k.

Consider now the side of the transformed net going from N_k to N_j. Of all the points in which this side intersects the circle about N_k, taken in their natural order on the side, there must be a last point P_1. Similarly, of all the points in which the side intersects the circle about N_j, taken in their natural order on the side, there must be a first point P_2. Form such pairs of points on all the sides of the network. Then there must be a minimum distance, 2δ, from the points on the segment of the network $P_1 P_2$ and the points on any other such segment, or the points on the H' circles about vertices other than N_k and N_j. Take $Q_1, Q_2, \cdots, Q_h$ any sequence of points on this segment P_1P_2 such that the distances $P_1 Q_1, Q_1 Q_2, \cdots, Q_h P_2$ are all less than δ. Replace the side $N_k N_j$ by the broken line $N_k P_1 Q_1 \cdots Q_h P_2 N_j$, and similarly replace every other side of a mesh of the net by such a broken line. Then each mesh of the net will go into a simply connected polygonal region, and no point of this region will lie outside of the original mesh, and abutting meshes. Since the distance from a point in s to a point in an abutting mesh is at most $2\sqrt{2}\eta < 3\eta$, if $\overline{T}_2$ is any continuous biunivocal transformation taking each mesh of the net on s into the corresponding polygonal mesh of the modified network just described, we shall have

$$\text{maximum } [T_1(p) - \overline{T}_2(p)] < 3H.$$

Since the inverse transformation $\overline{T}_2^{-1}$ takes a point in S into the same or an abutting mesh into which T_1^{-1} takes this point, we have

$$\text{maximum } [T_1^{-1}(P) - \overline{T}_2^{-1}(P)] < 3\eta < 3H.$$

We must now go from the polygonal network to convex polygons of the kind described in the lemma. We use T_1^{-1} to transform the vertices of the polygonal network back to s. This gives a series of points $n_k\ p_1\ q_1 \cdots q_h\ p_2\ n_j$ on the segment $n_k\ n_j$ dividing it into consecutive segments which we may put into correspondence with the segments of the broken line $N_k\ P_1\ Q_1$

$\cdots Q_h P_2 N_j$. We thus have a correspondence set up between the periphery of a rectangle in s and of a possibly reentrant polygon in S, for each mesh. We wish to show that it is possible to divide the two figures by straight lines into a finite number of similarly placed triangles.

Consider the polygon. We shall say that a point on its boundary is *visible* from another point situated anywhere in its plane, if the two can be joined by a straight line which does not cut the boundary. For any vertex of the polygon, we may find a point inside the polygon, sufficiently close to it so that this vertex and the two adjacent ones are visible from it. If the polygon is not a triangle, we may shift this point so that from it some fourth vertex, as well as as the three consecutive ones we began with, are visible.

Let us now consider one of the rectangles, and its corresponding, possibly reentrant, polygon. Select a vertex of the rectangle, and the corresponding vertex of the polygon. Find a point inside the polygon from which this vertex, the two adjacent ones, and a fourth are visible, and join it to these four by straight lines. Select a point inside the rectangle, so near the chosen vertex that, when the corresponding straight lines are drawn, all the new polygons will be *convex*. Now proceed similarly with one of the vertices of the convex polygons which is not collinear with its adjacent vertices. At each stage we reduce the number of sides of the polygon dealt with, and keep the new polygons convex, so as to allow the process to be repeated. After a finite number of repetitions, no polygon will have more than three sides, and we shall have succeeded in dividing our corresponding meshes into similarly placed triangles.

The sides of these triangles in the two networks will have a lower bound σ. Surround each node of each network with a circle of radius $\sigma/3$. Draw the inscribed polygon meeting the circle in the points where this circle cuts the lines of the network. These polygons will be convex, as will be polygons formed from the triangles by removing the portions inside these polygons. Further, the network of convex polygons will have the property that precisely three of them meet at each vertex inside the square, two at each vertex on a side of the square, and just one abuts on each vertex of the square.

Suppose, now, that T_2 is a continuous, biunivocal transformation which takes each convex polygon in s into its corresponding convex polygon in S. Points in S arising from a given point in s under T_2 and $\bar{T}_2$ respectively come from the same mesh or adjacent meshes under $\bar{T}_2$. Consequently, under T_1 these points come from meshes at worst next but one, or points in s at distance less that $3\sqrt{2}\,\eta < 5\eta$, and we shall have

$$\text{maximum } [T_1(p) - T_2(p)] < 5H < \epsilon/2.$$

Similarly, the inverse transformations, T_2^{-1} and T_1^{-1}, take a point in S into two points of s at worst in next but one meshes, and we have

$$\text{maximum } [T_1^{-1}(P)-T_2^{-1}(P)]<5\eta<5H<\epsilon/2.$$

Thus the network of convex polygons which we have found satisfies all the conditions of Lemma 1, since the distance between T_2 and T_1 is less than $\epsilon/2$.

IV. Differential transformations

We shall now show that a transformation T_2 of the type mentioned in Lemma 1 may be found which is defined by functions having partial derivatives with certain continuity properties. That is, we shall prove

Lemma 2. *Given two squares, s and S, subdivided into corresponding convex polygons, of which the number meeting at a vertex is one, two, or three according as the vertex is a vertex of the square, a point on a side of the square, or an interior point of the square, respectively; there exists a continuous biunivocal transformation T_2: $X = X_2(x,y)$, $Y = Y_2(x,y)$, for which X_2 and Y_2 possess first partial derivatives with respect to x and y continuous in s, and for which the jacobian of X_2, Y_2 with respect to x and y never vanishes in s, which maps the interior and boundary points of each convex polygon in s on those of the corresponding polygon in S.*

In setting up the differentiable transformation T_2, we shall frequently have occasion to change our system of coördinates. To carry over the properties of the transformation from one of these systems to another, we prove a general result once for all.

If a point transformation of one region on another is given parametrically, the parametric curves $U=U(X, Y)$, $V=V(X, Y)$, $u=u(x, y)$, $v=v(x, y)$ being such that (1) *the (U, V)-(X, Y) relation is one-to-one, so that it may be solved in the form $X=X(U, V)$, $Y=Y(U, V)$, and similarly the (u, v)-(x, y) relation may be solved in the form $x=x(u, v)$, $y=y(u, v)$;* (2) *the functions $U(X, Y)$ and $V(X, Y)$ possess continuous first partial derivatives with respect to X and Y, and similarly the functions $u(x, y)$ and $v(x, y)$ possess continuous first partial derivatives with respect to x and y;* (3) *the jacobian of U, V with respect to X, Y never vanishes, and similarly that of u, v with respect to x, y never vanishes; then a one-to-one relation between U, V and u, v: $U=U(u, v)$, $V=V(u, v)$, for which U and V possess continuous first partial derivatives with respect to u and v, and for which the jacobian of U, V with respect to u, v never vanishes, implies a one-to-one relation between X, Y and x, y: $X=X(x, y)$, $Y=Y(x, y)$, with similar properties at interior points of the regions, and conversely.*

For, by hypothesis (1), X and Y are single-valued functions of U and V. Also, by hypotheses (2) and (3), these functions possess continuous first partial derivatives with respect to U and V.* But, we are also assuming a differentiable one-to-one relation of U, V to u, v, and by hypotheses (1) and (2), there is a differentiable one-to-one relation between u, v and x, y. Since a differentiable, uniform function of a pair of similar functions yields a similar function, we see that X and Y are uniform, differentiable functions of x and y. The property of the jacobian follows from the fact that

$$\frac{\partial(X,Y)}{\partial(x,y)} = \frac{\partial(X,Y)}{\partial(U,V)} \frac{\partial(U,V)}{\partial(u,v)} \frac{\partial(u,v)}{\partial(x,y)},$$

and

$$1 = \frac{\partial(X,Y)}{\partial(U,V)} \frac{\partial(U,V)}{\partial(X,Y)}.$$

The second relation shows that the first factor above can not vanish, being the reciprocal of a finite quantity, and the second and third factors do not vanish by our hypothesis.

The converse is proved in a similar manner.

The principal types of coördinate systems we shall use are three. Those of the first type will be rectangular cartesian coördinates, with various orientations, and choice of origin. These clearly satisfy our three hypotheses, with relation to a fundamental coördinate system, say that with origin at one corner of the square s or S, and axes along two of its sides.

The second type are polar coördinates, with the pole at an arbitrary place. It is evident that these satisfy all our conditions, for any region not including the pole, when the angular coördinate t is reduced modulo 2π, and all the functions involving this coördinate are assumed to be of period 2π with respect to it.

The third type are modified polar coördinates. We begin with a closed curve which surrounds the pole, or origin, with a continuously turning tangent which never coincides with a radius vector. We use a series of curves, similar to this and similarly placed with reference to the pole, as our u-curves, taking the scale such that the length cut off by these curves on any fixed radius vector increases uniformly with u. The v curves are our radii vectores. Analytically, if $r=f(t)$ is the equation of the closed curve in polar coördinates, our parametric curves are related to polar coördinates by the equations

$$u=a+b\,r f(t), \quad v=t.$$

* Cf. Osgood, *Funktionentheorie*, Berlin, 1920, p. 69.

We may now show that, in any region not including the pole, these parametric curves satisfy our three hypotheses. In such a region, a pair of values x, y yields a single pair of values r, t and hence a single pair of values u, v. Conversely, a pair u, v fixes t by the second relation and then r by the first and thence x, y. Thus the first hypothesis is satisfied. The partial derivatives with respect to r and t are $\partial u/\partial r = bf(t)$, $\partial u/\partial t = br\,f'(t)$, $\partial v/\partial r = 0$, $\partial v/\partial t = 1$. These are all continuous, since our assumption about the tangent to the closed curve involves the existence of $f'(t)$, finite and continuous. From the relation of polar to cartesian coördinates, we see that, in a region excluding the pole, the second hypothesis is satisfied. For the third hypothesis, we have

$$\frac{\partial(u,v)}{\partial(x,y)} = \frac{\partial(u,v)}{\partial(r,t)}\,\frac{\partial(r,t)}{\partial(x,y)} = \frac{bf(t)}{r},$$

which does not vanish since $f(t)$ does not vanish.

We next proceed to the explicit construction of the differentiable transformation T_2. Our method is first to set up affine transformations applicable to suitable neighborhoods of the vertices of the polygonal network, then to interpolate between these along the sides of the polygons, and finally to interpolate into their interiors.

Since only one polygon abuts on the vertices of the squares s and S, at these points we may use the identity, or a similarity transformation. We use the latter:

$$X = kx, \quad Y = ky,$$

taking k positive and so small that the adjacent vertices of the network in s are mapped by this transformation on the adjacent sides of the network in S. At vertices of the network on a side of the square, we have to transform this side, and a line meeting it, into the corresponding side and a line meeting it. Taking the vertices in question as our origin, and the sides of the squares as the y axes, so oriented that, with the standard relation of the axes, the positive x axes point inside the squares, we have to take the lines $x=0$, $y=mx$ into $X=0$, $Y=MX$ respectively, where m and M may be positive or negative, by a transformation with positive jacobian. We may use

$$X = kx, \quad Y = ky + k(M - m)x,$$

where we again take k positive and so small that the three adjacent vertices of the network in s are each mapped on the interior of the corresponding adjacent side in S. At the interior vertices of the network, we have to transform three lines forming three convex sectors into three similar lines. Taking

the vertices in question as our origin, and one of the lines as the positive x axis, we have to transform the lines $y=0$, $y=mx$, $y=-nx$ into $Y=0$, $Y=MX$, $Y=-NX$ respectively, where m, n, M and N are positive, by a transformation with positive jacobian. We may use

$$X = kmn(M+N)x + k(mN - nM)y,$$

$$Y = kMN(m+n)y,$$

where we take k, as before, positive and so small that the three adjacent vertices of the network in s are each mapped on the interior of the corresponding adjacent side in S.

We now have affine transformations defined at each vertex of our network. We surround each vertex in s by a small triangle, with sides perpendicular to the sides of the network meeting there, and define T_2 inside and on the boundary of this triangle as the affine transformation just referred to. This gives a corresponding set of triangles in S. We must now interpolate between these transformations along the sides of the polygon. Let $p_1\, p_2$ be a side in s, and $P_1\, P_2$ be the corresponding side in S. If we take p_1 as origin, and $p_1\, p_2$ as the positive x axis, and P_1 as origin and $P_1\, P_2$ as the positive X axis, the affine transformation at p_1 will have as its equations

$$X = a_1x + b_1y, \quad Y = c_1y,$$

where a_1 is positive from our choice of direction along the x axis, and c_1 is positive since the jacobian remains positive for all axes with the standard orientation. In terms of the same coördinates, the affine transformation at p_2 will be

$$X = a_2x + b_2y + d_2, \quad Y = c_2y.$$

Here a_2 and c_2 are positive as before. We see that $d_2 = P_1\, P_2 - a_2(p_1\, p_2)$ is positive, since our condition on adjacent vertices makes the image of p_1 under this transformation go into a point between P_1 and P_2, i.e. $(d_2, 0)$, the image of $(0, 0)$, is a point with positive abscissa. We also assume that $a_2 \geqq a_1$, which is permissible, since, if it were not the case, we could reverse the rôles of p_1 and p_2 which would interchange these quantities.

If $x=x_1$ and $x=x_2$ are the equations of the sides of the small triangles which cut $p_1\, p_2$, our problem is to find a differentiable, biunivocal transformation which agrees up to the first partial derivatives with the transformation at p_1 for $x=x_1$ and with that at p_2 for $x=x_2$. We write

$$
\begin{aligned}
X = {} & a_1\frac{(x-x_2)^2}{(x_1-x_2)^2}\left\{x_1+\frac{(x_1+x_2)(x-x_1)}{x_2-x_1}\right\} \\
& + a_2\frac{(x-x_1)^2}{(x_2-x_1)^2}\left\{x_2+\frac{(x_1+x_2)(x-x_2)}{x_1-x_2}\right\} \\
& + b_1y\frac{(x-x_2)^2}{(x_1-x_2)^2}\left\{1+\frac{2(x-x_1)}{x_2-x_1}\right\} \\
& + b_2y\frac{(x-x_1)^2}{(x_2-x_1)^2}\left\{1+\frac{2(x-x_2)}{x_1-x_2}\right\} \\
& + d_2\ \frac{(x-x_1)^2}{(x_2-x_1)^2}\left\{1+\frac{2(x-x_2)}{x_1-x_2}\right\}, \\
Y = {} & c_1y\frac{(x-x_2)^2}{(x_1-x_2)^2}\left\{1+\frac{2(x-x_1)}{x_2-x_1}\right\} \\
& + c_2y\frac{(x-x_1)^2}{(x_2-x_1)^2}\left\{1+\frac{2(x-x_2)}{x_1-x_2}\right\}.
\end{aligned}
$$

By direct calculation, we find that, for $x=x_1$, these equations give

$$
\begin{aligned}
&X = a_1x_1+b_1y, && \partial X/\partial x = a_1, && \partial X/\partial y = b_1, \\
&Y = c_1y, && \partial Y/\partial x = 0, && \partial Y/\partial y = c_1.
\end{aligned}
$$

For $x=x_2$, they give

$$
\begin{aligned}
&X = a_2x_2+b_2y+d_2, && \partial X/\partial x = a_2, && \partial X/\partial y = b_2, \\
&Y = c_2y, && \partial Y/\partial x = 0, && \partial Y/\partial y = c_2.
\end{aligned}
$$

Thus this transformation joins to those at p_1 and p_2 as desired. We also note that, for $y=0$, we have

$$
\begin{aligned}
\partial X/\partial x = {} & a_1\frac{(x_2-x)}{(x_2-x_1)^3}\left\{(x_2-x_1)^2-3(x_1+x_2)(x-x_1)\right\} \\
& + a_2\frac{(x-x_1)}{(x_2-x_1)^3}\left\{(x_2-x_1)^2+3(x_1+x_2)(x_2-x)\right\} \\
& + d_2\frac{6(x-x_1)(x_2-x)}{(x_2-x_1)^3}.
\end{aligned}
$$

For $0<x_1<x<x_2$, the coefficients of a_2 and d_2 are positive. Since the sum of the coefficients of a_1 and a_2 is unity, and we have arranged matters so that $a_2 \geqq a_1 > 0$, and $d_2>0$, we see that the sum of the three terms is positive.

Hence, along the line $y=0$, X is a monotonic function of x, and the mapping is one-to-one. Again, for $y=0$, we have

$$\partial Y/\partial x = 0,$$

$$\partial Y/\partial y = c_1 \frac{(x-x_2)^2}{(x_1-x_2)^2}\left\{1+\frac{2(x-x_1)}{x_2-x_1}\right\} + c_2\frac{(x-x_1)^2}{(x_2-x_1)^2}\left\{1+\frac{2(x-x_2)}{x_1-x_2}\right\}.$$

For $0<x_1<x<x_2$, the coefficients of c_1 and c_2 are positive, and since c_1 and c_2 are positive, $\partial Y/\partial y$ is positive. Hence the jacobian of X, Y with respect to x, y for the cubic transformation under discussion is positive along the line $y=0$. Thus the transformation is one-to-one im kleinen, and, since it is one-to-one along the line $y=0$, and continuous, we may find a strip including this line in which it is one-to-one im grossen. Furthermore, since the transformation has continuous derivatives, if this strip is taken sufficiently small, the jacobian of the transformation will be positive throughout the strip.

We have now to interpolate into the interior of our polygons. In each polygon of the figure s, we draw a differentiable closed curve, close to the boundary of the polygon. We may form it of straight lines parallel to the sides between the x_1 and x_2 used above in the strips, and circular arcs in the triangular sectors. If it is taken sufficiently close to the polygon, it will be visible from some point c inside the polygon, i.e. such that each radius vector drawn from c to the curve cuts it in a single point, at an angle not zero. Thus it may be used as the closed curve of a system of modified polar coördinates u, v of the kind previously described, with the pole at c. Consider the transforms of these u curves in S, under the cubic and affine transformations. In a sufficiently narrow strip, these curves will be visible from some point C inside the polygon. In this strip we shall take as parametric curves the curves $U=\text{const.}$, the transforms of $u=\text{const.}$, and the curves $T=\text{const.}$, the radii vectores drawn from C, T being the angle made with a fixed vector. To show that these are admissible parameters, we note that the one-to-one relation of U, T to X, Y follows from the fact that the U curves are visible from C, so that a pair of values U, T gives a single point X, Y. On the other hand, a pair X, Y clearly gives one value of T, and a single U resulting from the u obtained from x, y. The U and T curves clearly are differentiable. For the condition on the jacobian, we have

$$\frac{\partial(U,T)}{\partial(X,Y)} = \frac{\partial(U,T)}{\partial(U,V)}\,\frac{\partial(U,V)}{\partial(x,y)}\,\frac{\partial(x,y)}{\partial(X,Y)},$$

where the curves $V=\text{const.}$ are the transforms of $v=\text{const.}$ under the affine and cubic transformations. From the definition of the U, V curves, the

jacobian of U, V with respect to x, y is the same as that of u, v with respect to x, y and hence can not vanish, since u, v were admissible parameters. The last jacobian does not vanish, since it relates to the cubic and affine transformations. For the first factor we have

$$\frac{\partial(U,T)}{\partial(U,V)} = \frac{\partial T}{\partial V}.$$

If this vanished, we should either have $\partial T/\partial U = 0$ at the same time, in which case since U and V are admissible parameters, we would be at a singular point on the curve (or in its representation), or we should have $\partial T/\partial U \neq 0$, and the T and U curves would be tangent. The first case is inadmissible, since the angular coördinate represents the radii without singularities in any admissible system of coördinates, and the second is excluded since the U curves we are using are all visible from C. Thus the U, T parameters satisfy our three conditions, and form an admissible system.

We choose the scales so that the curves $u=0$ and $u=1$ are respectively the inner and outer boundaries of a strip entirely inside the region where our affine and cubic transformations are valid, while their transforms $U=0$ and $U=1$ are the boundaries of a strip inside which the U, T coördinates are admissible. In these coördinates, the affine and cubic transformations will have the equations

$$u = U, \quad t = F(T,U),$$

where we replace v by t, its equal. We note that

$$\frac{\partial F}{\partial T} = \frac{\partial(u,t)}{\partial(U,T)} > 0.$$

In the strip between $u=1$ and the polygon, we define T_2 as the affine and cubic transformation, which is that written above near $u=1$. In the strip from $u=1$ to $u=0$, we use a derived transformation, such that it agrees with this transformation up to the first derivatives at $u=1$, and takes the curves $t=$const. into curves *tangent to the curves* $T=$const. at $u=0$. The derived transformation is

$$u = U, \quad t = F(T,2U^2 - U^3) = F_1(T,U).$$

This is one-to-one for values of u, or U, between 0 and 1. For, since $u=U$, $t=F(T, U)$ is biunivocal for these values, for any value of u or U in this range, $t=F(T, U)$ gives a one-to-one relation of t to T. But $2U^2-U^3$ takes on the values between 0 and 1 when U does, and accordingly for fixed u or U, $t=F_1(T, U)$ gives a one-to-one relation of t to T. The tangency prop-

erty required follows from the fact that $(\partial F_1/\partial U)_{U=0}=0$, while $(\partial F_1/\partial T)_{U=0} =(\partial F/\partial T)_{U=0}>0$. For $u=U=1$, we easily see that the derived transformation and its derivatives agree with the earlier transformation. Its jacobian is $\partial F_1/\partial T=\partial F/\partial T>0$.

In continuing our interpolation, it will be convenient to use polar coördinates. As our lemma on systems of coördinates only applies to interior points, we must extend our transformation beyond the inner boundary, $u=U=0$, to make this curve an interior curve. We do this by putting

$$u = U, \quad t = F_1(T,0),$$

in a narrow strip inside the curves $u=U=0$. This joins with our preceding transformation along these curves, continuity of the derivatives being retained, since $(\partial F_1/\partial U)_{U=0}=0$. Its jacobian is $(\partial F_1/\partial T)_{U=0}>0$.

We next draw two circles of radius a about the points c and C as centers respectively, taking a so small that these circles lie entirely inside the curves $u=0$ and $U=0$. To extend our transformation so that it maps the region between $u=0$ and the first circle on to that between $U=0$ and the second, we introduce polar coördinates t, r and T, R. Let the equations of $u=0$ and $U=0$ in these coördinates be

$$r = w(t) = v(T), \quad R = V(T),$$

where, as in the following equations, all the functions of t and T are of period 2π. Along these curves we have

$$\partial t/\partial R = 0, \quad \partial t/\partial T = (\partial F_1/\partial T)_{U=0}, \quad \partial r/\partial R = E(T), \quad \partial r/\partial T = G(T).$$

Since $dr/dT=(\partial r/\partial R)\ (dR/dT)+\partial r/\partial T$, we have $v'(T)=E(T)V'(T)+G(T)$. We put

$$t = F_1(T,0), \quad r = G(R,T),$$

where $G(R, T)$ is periodic in T, has $\partial G/\partial R>0$, and satisfies the further conditions

$$G(a,T) = a, \quad (\partial G(R,T)/\partial R)_{R=a} = 1,$$
$$G(V,T) = v, \quad (\partial G(R,T)/\partial R)_{R=V} = E(T).$$

The determination of $G(R, T)$ for a fixed value of T reduces to the problem of finding a function of R, with positive derivative, which, with its derivative takes two positive assigned values at each end of an interval. The function $E(T)$ is positive, since

$$\left(\frac{\partial(r,t)}{\partial(R,T)}\right)_{U=0} = E(T)\left(\frac{\partial F_1}{\partial T}\right)_{U=0} > 0, \text{ and } \left(\frac{\partial F_1}{\partial T}\right)_{U=0} > 0.$$

We define the auxiliary function $H(R)$ as follows:

$$H(R) = 1 + (R - a)(M - 1)/k, \qquad a \leqq R \leqq a + k,$$
$$H(R) = M, \qquad a + k \leqq R \leqq V - k,$$
$$H(R) = E + (V - R)(M - E)/k, \; V - k \leqq R \leqq V.$$

Here we take

$$M = \frac{v - a - \frac{1}{2}(1 + E)k}{V - a - k},$$

and k a positive number so small that

$$k < \tfrac{1}{2}(V-a), \quad k < \frac{2(v-a)}{1+E}.$$

This makes $H(R)$ a continuous, broken-line function, with positive ordinate throughout, and such that the area under it is $v-a$. Also, $H(a)=1$ and $H(V)=E$. Consequently if we put

$$G(R,T) = a + \int_a^R H(R)dR,$$

it will satisfy all the conditions required of it above. We may take k an absolute constant, since the right members of the inequalities it satisfies are positive continuous periodic functions of T, and hence have positive minima.

We note that the transformation given above maps our ring shaped regions on one another in a one-to-one manner, since the relation $t=F_1(T,0)$ is clearly one-to-one, and if either t or T is fixed, the second relation $r=G(R, T)$ is one-to-one, since $\partial G/\partial R>0$. On the outer boundary it agrees with our former transformation, since $\partial r/\partial R = E(T)$ there, and $t=F_1(T, 0)$, $r=G(V, T)=v(T)$. From these we deduce

$$dr/dT = (\partial r/\partial R)(dR/dT) + \partial r/\partial T, \text{ or } v'(T) = E(T)V'(T) + \partial r/\partial T,$$

which, combined with the earlier relation $v'(T)=E(T)V'(T)+G(T)$, shows that $\partial r/\partial T=G(T)$. The transformation has a positive jacobian, since this equals $(\partial F_1/\partial T)_{U=0}(\partial G/\partial R)$.

In the ring between $R=a$ and $R=a/2$, corresponding to $r=a$, $r=a/2$, we put

$$r = R,$$
$$t = 4T(R - a)^2(4R - a)/a^3 + F_1(T,0)(2R - a)^2(5a - 4R)/a^3.$$

For $R=a$, this gives us $r=a$, $\partial r/\partial R=1$, $\partial r/\partial T=0$, $t=F_1(T, 0)$, $\partial t/\partial T=(\partial F_1/\partial T)_{U=0}$, $\partial t/\partial R=0$. These values agree with those of the last trans-

formation at this circle, all except the value of $\partial r/\partial R$ following directly, and that from the relation

$$dr/dT = (\partial r/\partial R)(dR/dT) + \partial r/\partial T, \quad \text{or} \quad 0 = \partial r/\partial T.$$

For $R=a/2$, the equations give $r=a/2$, $\partial r/\partial R=1$, $\partial r/\partial T=0$, $t=T$, $\partial t/\partial T=1$, $\partial t/\partial R=0$. Thus it joins on to the identity along this circle. The transformation is one-to-one, since, for fixed r or R, the t-to-T relation is one-to-one. For, we have

$$\frac{\partial t}{\partial T} = \frac{4(R-a)^2(4R-a)}{a^3} + \left(\frac{\partial F_1}{\partial T}\right)_{U=0} \frac{(2R-a)^2(5a-4R)}{a^3},$$

which is clearly positive for $a/2<R<a$. As this is the value of the jacobian, we see that our transformation has a positive jacobian.

Inside the circles $r=a/2$, $R=a/2$, we use the identity transformation

$$r = R, \quad t = T,$$

to complete the mapping. As the identity has the same form in all coördinates, the singularity of the coördinates at the pole is of no importance.

There is now a transformation for each polygon, given in polar coördinates with pole in the polygon, and defined in a region including the polygon in its interior. Thus, by our lemma on systems of coördinates, it may be expressed in terms of the particular set of cartesian coördinates used for T_2. The totality of these yields a differentiable transformation T_2 of the kind required.

V. Analytic transformations

Weierstrass* has given an example of an analytic function approximating to any degree of accuracy a given continuous function of two variables, $f(x, y)$, in a finite region R. This approximating function is

$$f_k(x,y) = \frac{1}{k^2\pi}\int_R\int f(X,Y)e^{-((X-x)/k)^2-((Y-y)/k)^2}\,dXdY,$$

when k is sufficiently large. It is easy to show that if $f(x, y)$ possesses continuous first partial derivatives over any closed region D interior to R, we have

$$\lim_{k\to\infty} f_k(x,y) = f(x,y)\ ; \ \lim_{k\to\infty}(\partial f_k/\partial x) = \partial f/\partial x\ ; \ \lim_{k\to\infty}(\partial f_k/\partial y) = \partial f/\partial y$$

uniformly over D.

---

* Weierstrass, Berliner Sitzungsberichte, 1885.

This result enables us to approximate to any biunivocal, differentiable transformation by an analytic transformation, the approximation applying to the derivatives, but the analytic transformation thus obtained is not necessarily one-to-one. For the case in which the differentiable transformation has a non-vanishing jacobian, we may show that, when the approximation is sufficiently close, it is biunivocal, and shall prove the following lemma:

LEMMA 3. *If*

$$X = X(x,y),\ Y = Y(x,y)$$

is a continuous, biunivocal transformation, holding in the closed convex region s, for which X and Y possess first partial derivatives with respect to x and y, continuous in s, and for which the jacobian of X and Y with respect to x and y never vanishes in s, and

$$X = X_a(x,y),\ Y = Y_a(x,y)$$

are a pair of differentiable functions, which approximate X and Y, and whose first partial derivatives approximate those of X and Y, uniformly in s, then, if the approximation is sufficiently close, these equations represent a biunivocal transformation for the interior of s.

We note from the form of the approximating functions, that they take each point x, y into a point X_a, Y_a, but may take two points x, y into the same point X_a, Y_a. We must show this to be impossible when the approximation is sufficiently good. We shall divide the proof into two parts, first proving the lemma for the case in which the given functions X, Y are linear functions, and then for the general case.

1. Let

$$X = ax + by + e,\ Y = cx + dy + f$$

be the given transformation. To predict the necessary degree of approximation, we make some preliminary calculations. Consider the radius drawn out from any fixed point p_1 to a variable point p, which moves out along this radius in s. If the line makes an angle A with the X axis, we shall have

$$X - X_1 = a(x - x_1) + b(y - y_1) = r(a\cos A + b\sin A),$$
$$Y - Y_1 = c(x - x_1) + d(y - y_1) = r(c\cos A + d\sin A).$$

For the distance in the transformed figure, S, we have

$$R^2 = (X - X_1)^2 + (Y - Y_1)^2 = r^2[(a^2 + c^2)\cos^2 A + 2(ab + cd)\cos A\sin A + (b^2 + d^2)\sin^2 A].$$

As A varies, the ratio R^2/r^2 varies between the maximum and minimum values

$$\tfrac{1}{2}(a^2 + b^2 + c^2 + d^2) \pm \tfrac{1}{2}\sqrt{(a^2 + c^2 - b^2 - d^2)^2 + 4(ab + cd)^2}.$$

The product of these is $(ad-bc)^2$, which is the square of the jacobian. This might have been predicted since the square root of the product of the maximum and minimum values, and the jacobian both measure the ratio of the area of a circle in s to that of its corresponding ellipse in S. As the jacobian does not vanish, neither the maximum nor minimum value is zero. Let the value of the minimum be q^2. Then, for the affine transformation, the minimum value of R/r is $q>0$.

Now consider the image of the variable radius $p_1\,p$ under the approximating transformation. Let the transformed radius be $P_1\,P$, where $P_1=(X_1, Y_1)$ and $P=(X, Y)$. We have for its length

$$R = \sqrt{(X - X_1)^2 + (Y - Y_1)^2}.$$

We calculate its derivative with respect to r, and find

$$\frac{dR}{dr} = \frac{[(X-X_1)(\partial X/\partial x)+(Y-Y_1)(\partial Y/\partial x)]\cos A}{\sqrt{(X-X_1)^2+(Y-Y_1)^2}} + \frac{[(X-X_1)(\partial X/\partial y)+(Y-Y_1)(\partial Y/\partial y)]\sin A}{\sqrt{(X-X_1)^2+(Y-Y_1)^2}}.$$

We may write this in a form involving derivatives only, by noting that $(X-X_1)$ and $(Y-Y_1)$ enter, essentially, through their ratio, only. But, since the image of the straight line $p_1\,p$ is a differentiable curve, having $P_1\,P$ as a chord, there is some intermediate point on this curve for which the tangent is parallel to the chord. If this point is $\overline{P}$ we shall have

$$\frac{Y - Y_1}{X - X_1} = \frac{d\overline{Y}}{d\overline{X}} = \frac{(\partial \overline{Y}/\partial x)\cos A + (\partial \overline{Y}/\partial y)\sin A}{(\partial \overline{X}/\partial x)\cos A + (\partial \overline{X}/\partial y)\sin A}.$$

We notice that both terms of this ratio can not be zero unless the jacobian of the transformation is zero at $\overline{P}$. We may now write for the derivative

$$\frac{dR}{dr} = \frac{(X_x\cos A+X_y\sin A)(\overline{X}_x\cos A+\overline{X}_y\sin A)}{\sqrt{(\overline{X}_x\cos A+\overline{X}_y\sin A)^2+(\overline{Y}_x\cos A+\overline{Y}_y\sin A)^2}} + \frac{(Y_x\cos A+Y_y\sin A)(\overline{Y}_x\cos A+\overline{Y}_y\sin A)}{\sqrt{(\overline{X}_x\cos A+\overline{X}_y\sin A)^2+(\overline{Y}_x\cos A+\overline{Y}_y\sin A)^2}}$$

where the subscripts indicate partial derivatives. If in this expression we replace the partial derivatives X_x, X_y, Y_x, Y_y, as well as $\overline{X}_x$, $\overline{X}_y$, $\overline{Y}_x$, $\overline{Y}_y$,

by a, b, c, d, their values for the affine transformation, it reduces to R/r for the affine transformation, or

$$\sqrt{(a^2 + c^2)\cos^2 A + 2(ab + cd)\cos A \sin A + (b^2 + d^2)\sin^2 A} \geqq q > 0.$$

As the expression is a continuous function of these partial derivatives for values in some neighborhood of a, b, c, d, a u_A may be found such that, if these partial derivatives are all within u_A of their values for the affine transformation, the expression will be $>q/2$, for a given A. But, as the function is continuous in A, for values in some neighborhood of this given value, it will be $>q/4$. By the Heine-Borel theorem, a finite number of such neighborhoods may be found which include all values of A, and hence a number u may be found less than any of the corresponding u_A. If the approximation has its partial derivatives within u of those of the affine transformation, for any A, we shall have $dR/dr>q/4>0$.

We now assert that a transformation approximating the affine transformation to this degree of accuracy is necessarily one-to-one in s. For, if it were not, it would take two points in s, p_2 and p_3, into the same point P_2. Since s is convex, $p_2 p_3$ lies entirely in s. Let p_1 be a point on the extended line $p_3 p_2$, and P_1 its transform. Let a variable point p, at distance r from p_1, move along $p_2 p_3$, and let the distance of its transform, P, from P_1 be R. As r changes from $p_1 p_2$ to p_1p_3, R changes from $P_1 P_2$ back to P_1P_2. Hence, at some point between p_2 and p_3, we must have $dR/dr=0$. But this is impossible, since $dR/dr>q/4>0$.

This proves the lemma for the affine case. We notice that here the limit on the approximation involves only the derivatives.

2. Now consider a general differentiable transformation, with non-vanishing jacobian in a closed region s. At every point p in this region, we form an approximating linear transformation, having the same derivatives as the given transformation has at this point. By the first part of this proof, we may find a u_p for this linear transformation such that any transformation whose derivatives approximate its derivatives to within u_p will be one-to-one. Since the given transformation has continuous derivatives, we may surround each point by a circle of radius v_p such that, in this circle, the partial derivatives never differ from those at the center by more than $u_p/2$. Surround each point p by a circle of radius $v_p/2$. By the Heine-Borel theorem, we may find a finite number of such circles such that each point of s is inside some one of them. Let u be the minimum u_p for any of these points, and v the minimum v_p. We see that, if any point in s is taken as the center of a circle of radius $v/2$, the partial derivatives of the given transformation will not

differ by more than $u/2$ from some linear transformation for which u may be taken as the u of part one.

Since the given transformation is continuous in s, its inverse is, and from the nature of s we may find a V such that, if the images of two points are at distance less than V, the points themselves will be at distance less than $v/4$. We now assert that any approximating transformation which takes a point in s into a point at distance less than V from its transform under the given transformation, and whose derivatives are within $u/2$ of those of the given transformation, is necessarily one-to-one. For, suppose it took two points p_1 and p_2 into the same point P'. Let the transforms of these points under the given transformation be P_1 and P_2. From the nature of our approximation, P_1P' and P_2P' are both less than V. Hence P_1P_2 is less than $2V$, and $p_1\,p_2$ is less than $v/2$. Draw a circle with center p_1 and radius $v/2$. If this circle does not lie entirely in s, the part of it contained in s is convex, since s is convex, and hence this part may be used as the region of part one. But, in this circle, the derivatives of the given transformation do not differ by more than $u/2$ from those of some linear transformation for which this is the u of part one. But, since the derivatives of the approximating transformation are within $u/2$ of those of the given transformation, they are within u of those of this linear transformation. Thus the approximating transformation is one-to-one in this circle, and can not take p_1 and p_2 into the same point. Thus our lemma is proved.

We are now in a position to prove

THEOREM I. *Given a pair of continuous functions*

$$X = X(x,y), \quad Y = Y(x,y),$$

defining a continuous, biunivocal transformation of some closed, two-dimensional region r of finite connectivity of the x, y plane, into a closed region R of the X, Y plane, and a positive constant ϵ, there exists a pair of analytic functions

$$X = X_a(x,y), \quad Y = Y_a(x,y),$$

defining a biunivocal transformation of some closed region of the x, y plane including r, into a closed region of the X, Y plane including R, whose distance from the given transformation is less than ϵ.

For, from the discussion of §2 we may embed the regions r and R in two squares s and S, and find a continuous transformation T_1 of s on S, which agrees with the given transformation in r, and takes the vertices of s into those of S. Then we dissect the squares into convex polygons, by Lemma 1, such that any continuous biunivocal transformation T_2 which maps

each polygon of s into the corresponding polygon of S is at distance less that $\epsilon/2$ from T_1. By Lemma 2, we then find a pair of functions

$$X = X_2(x,y), \quad Y = Y_2(x,y),$$

with continuous first partial derivatives, whose jacobian never vanishes in s, and which has the properties demanded of T_2. Consequently its distance from T_1 is less than $\epsilon/2$. We next apply Lemma 3 to this transformation to find a V and a $u/2$ such that, if a pair of differentiable functions are given which approximate X and Y to within $V/\sqrt{2}$, and whose derivatives approximate those of X and Y to within $u/2$, they define a one-to-one transformation in s. Also we find a W such that the transforms of any two points in S whose distance is less than W, under the inverse of T_2, are at distance less than $\epsilon/2$. Finally, we use the method of Weierstrass recalled at the beginning of this section to find a pair of analytic functions

$$X = X_3(x,y), \quad Y = Y_3(x,y),$$

which approximate $X_2(x, y)$ and $Y_2(x, y)$ to within $\epsilon/2\sqrt{2}$, $V/\sqrt{2}$, $W/\sqrt{2}$, and whose derivatives approximate those of these functions to within $u/2$. These equations define the analytic transformation whose existence the theorem asserts. For, by Lemma 3, it is one-to-one in s. From the choice of the approximation, and of W, we have

$$[T_3(p) - T_2(p)] < \epsilon/2,$$
$$[T_3^{-1}(P) - T_2^{-1}(P)] = [T_2^{-1}T_2T_3^{-1}(P) - T_2^{-1}T_3T_3^{-1}(P)] < \epsilon/2.$$

Thus the distance of the transformation T_3 to T_2 is less than $\epsilon/2$, and since that of T_2 to T_1 is less than $\epsilon/2$, the distance from T_3 to T_1 is less than ϵ, in s. Since no point of R is within ϵ of the boundary of S, T_3 maps s on a region including R, and hence T_3^{-1} is defined in a region including R. As the distance from T_3 to T_1 is less than ϵ wherever both are defined, we see that T_3 satisfies all the conditions demanded of it in the theorem.

COROLLARY. *The functions $X_3(x, y)$ and $Y_3(x, y)$ of the theorem may be taken as a pair of polynomials.*

For we may approximate to any analytic function, as well as to its first derivatives, by a polynomial.

VI. TRANSFORMATIONS ON A SPHERE

Our argument may be extended to transformations of a sphere into itself with but slight modification. A natural definition of an analytic transformation on a sphere is *a transformation which is expressed by analytic functions*

in terms of some standard set of coördinates, such as the polar coördinates ϕ and θ. To avoid trouble at the poles, we require that the functions be analytic and single-valued in ϕ and θ for *every* possible choice of the axis. In our discussion it will be more convenient to use an equivalent form of this definition, and note that any transformation of a three-dimensional region, not necessarily simply connected, including the points of the sphere as interior points, analytic in terms of cartesian coördinates, which leaves the sphere invariant, determines a transformation on the sphere which is analytic in the above defined sense, and conversely any transformation analytic in the earlier sense may be extended to give an analytic space transformation of the kind here used. With this definition, we may formulate

THEOREM II. *Given a set of three continuous functions*

$$X = X(x,y,z),\ Y = Y(x,y,z),\ Z = Z(x,y,z),$$

which leave the unit sphere invariant,

$$X^2 + Y^2 + Z^2 = 1 \text{ if } x^2 + y^2 + z^2 = 1,$$

and define a continuous biunivocal transformation on this sphere, and a positive number ϵ, *there exists a set of three analytic functions*

$$X = X_a(x,y,z),\ Y = Y_a(x,y,z),\ Z = Z_a(x,y,z),$$

which leave the unit sphere invariant,

$$X_a^2 + Y_a^2 + Z_a^2 = 1 \text{ if } x^2 + y^2 + z^2 = 1,$$

and define a biunivocal transformation of the unit sphere into itself whose distance from the given transformation is less than ϵ.

We divide the sphere by two networks of convex polygons, whose sides are great circles, the number of polygons meeting at each vertex being three, which have the property that any continuous biunivocal transformation T_2 of the sphere into itself, which maps each polygon of the first network on the corresponding polygon of the second network, is at distance less than $\epsilon/2$ from the given transformation, T_1. This is shown to be possible by the method used to prove Lemma I. To set up a differentiable transformation T_2, we begin with the neighborhood of a vertex and its transform, gnomonically project these neighborhoods on the tangent planes at the vertices in question, and set up affine transformations in these planes as in the proof of Lemma 2. They project back into differentiable transformations on the sphere, which will yield affine transformations when projected from the center on any tangent plane. To interpolate along a side, we work in the

tangent plane at its midpoint, and in that at the midpoint of the corresponding side, and form the cubic transformations of Lemma 2. These project into differentiable transformations on the sphere. For the interpolation into the interior of a polygon, we work in the tangent plane at some interior point of it, and that at some interior point of its transform, and proceed as in Lemma 2. We thus obtain a differentiable transformation of the sphere into itself which is one-to-one, possesses a non-vanishing jacobian with respect to any system of ϕ, θ coördinates, and approximating T_1 to within $\epsilon/2$. We readily extend this to a region between two spheres concentric with the unit sphere, and enclosing it, by adding to the ϕ, θ transformation the equation $R=r$. This transformation, when expressed in cartesian coördinates, gives three differentiable functions

$$X = X_2(x,y,z),\ Y = Y_2(x,y,z),\ Z = Z_2(x,y,z),$$

such that $X_1^2+Y_1^2+Z_1^2=1$ if $x^2+y^2+z^2=1$.

We shall now use the considerations of Lemma 3 to show that any differentiable transformation in three variables which approximates this one, both as to coördinates and their derivatives, sufficiently closely, will, when the transforms of points on the unit sphere under it are projected gnomonically on the sphere, yield a one-to-one transformation of the sphere. At every point p of the sphere we draw the tangent plane, and project the points in some neighborhood of p on this tangent plane. Similarly we project their transforms on the tangent plane at the transformed point P. This gives a differentiable transformation of a part of one of these planes on the other. We form an approximating linear transformation, having the same derivatives as this transformation at p. By part 1 of Lemma 3, we may find a u_p for this linear transformation such that any transformation of the plane whose derivatives approximate its derivatives to within u_p will be one-to-one. We surround each point p by a circle of radius v_p such that, in the circle which is the projection of this on the tangent plane, the partial derivatives of the plane transformation never differ from those at the center by more than $u_p/2$. We then surround each point p on the sphere by a circle of radius $v_p/2$. We pick out a finite number of such circles with centers at p_i such that each point on the sphere is inside some one of them. If, now, u is the minimum u_p and v the minimum v_p for the corresponding points, if any point q on the sphere is taken as the center of a circle of radius $v/2$, there is a point p_i on the sphere within a distance $v/2$ of q, such that if the points inside the circle at q are projected on the tangent plane to the sphere at p, and their transforms are projected on the tangent plane at P_i, the transform of p_i, a plane differentiable transformation will arise with the property that any

other differentiable transformation whose partial derivatives differ from those for it by less than $u/2$ will be one-to-one.

We now find a V for the transformation on the sphere such that, if the images of two points are at distance less than V the points themselves will be at distance less than $v/4$. Now consider any approximating differentiable transformation, T_3. Project the points in a circle of radius $v/2$ about q on the tangent plane at the nearest p_i, and their transforms under T_3 on the tangent plane at P_i, the transform of p_i under the given transformation. A plane differentiable transformation will arise whose partial derivatives are seen to be continuous functions of the partial derivatives of T_3, and certain quantities connected with the given transformation and its derivatives at p_i. Consequently we may find a u' such that, if the derivatives of T_3 are within u' of those of the given transformation, the derivatives of the plane transformation just found will be within $u/2$ of the plane transformation previously found for p_i. Thus the plane transformation will be one-to-one inside the circle of radius $v/2$. If T_3 in addition approximates the given transformation to within V in distance, it will project on the sphere into a transformation one-to-one throughout. For, if it took two points p_1 and p_2 into the same point P', whose transforms under the given transformation were P_1 and P_2, we should have P_1P' and P_2P' less than V, and hence P_1P_2 less than $2V$. Thus p_1p_2 would be less than $v/2$, and a circle with center at p_1 and radius $v/2$ could be drawn in which the projected transformation on the tangent plane at the nearest p_i would be one-to-one, contradicting the assumption that p_1 and p_2 gave a single point P'.

We now find, by the Weierstrass method quoted in Section 5, applied to functions of three variables, a set of three analytic functions $X_3(x, y, z)$, $Y_3(x, y, z)$, $Z_3(x, y, z)$ approximating the functions defining the differentiable transformation, $X_2(x, y, z)$, $Y_2(x, y, z)$, $Z_2(x, y, z)$, to within $\epsilon/4\sqrt{3}$, $V/\sqrt{3}$, $W/\sqrt{3}$, and whose derivatives approximate those of these functions to within u', where W is chosen such that two points of the sphere whose distance is less than W, under the inverse of the differentiable transformation, go into two points whose distance is less than $\epsilon/4$. We write, finally,

$$X = \frac{X_3(x,y,z)}{\sqrt{X_3^2 + Y_3^2 + Z_3^2}},$$

$$Y = \frac{Y_3(x,y,z)}{\sqrt{X_3^2 + Y_3^2 + Z_3^2}},$$

$$Z = \frac{Z_3(x,y,z)}{\sqrt{X_3^2 + Y_3^2 + Z_3^2}},$$

as the approximating analytic transformation. It clearly takes the sphere into itself, is one-to-one, and is at distance less than ϵ from the given transformation T_1, since it is at distance less than $\epsilon/2$ from T_2.

COROLLARY. *Given a continuous, biunivocal transformation of any simply connected closed surface into itself, a positive number* ϵ *and any set of curves on the surface deformable into meridians and parallels, we may find an analytic transformation, the analyticity being given in terms of the given curves, whose distance from the given transformation, measured in the same terms, is less than* ϵ.

For we have merely to deform the closed surface into a sphere, and the curves into meridians and parallels, and apply Theorem II.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.

ERRATA, VOLUME 28

M. H. INGRAHAM, *Solution of certain functional equations relative to a general linear set.*

Page 287, insert (1) at left of second displayed formula;

Page 289, formula (i) should read $v.).v^{0H}=1$;
last line above footnote should read

$$ {}_{h_1}C_h \equiv \prod_q {}_{h_1(\ )}C_{h(q)} \qquad (h_1,h) ; $$

Page 292, in right member of second displayed equation above the summation sign, for "$h_1 \pm h$" read "$h_1 \neq h$."

Page 293, in the first line of the proof of Theorem 2, after "distinct" insert "determinations of coefficients for";
fourth line from bottom should read

$$ = \sum_h v_h \sum_{h_1} {}_hC_{h_1} v_0^{h-h_1} r^{w(h_1)} v_1^{h_1} ; $$

third line from bottom should read

$$ = \sum {}_h v_h \sum_{k_1} \sum_{h_1}^{w(h_1)=k_1(p-1)} {}_hC_{h_1} v_0^{h-h_1} r^{w(h_1)} v^{h_1}; $$

last line, above second summation sign for "$w(h_1)\equiv(p-1)$" read "$w(h_1)\equiv k(p-1)$."

T. H. GRONWALL, *On the zeros of the function $\beta(z)$ associated with the gamma function.*

Page 395, line next to the bottom (footnote) for "$[f(z)-g(z)]g(z)$" read "$[f(z)-g(z)]/g(z)$";

Page 397, line 13, for "$1+O((\log n)/n)^2$" read

$$ O\left(\left(\frac{\log n}{n}\right)^2\right)." $$